United States
Department of
Agriculture

**Agricultural
Research
Service**

ARS-75

December 1988

Microcomputer Program for Daily Weather Simulation

ABSTRACT

Woolhiser, D.A., C.L. Hanson, and C.W. Richardson. 1988. Microcomputer Program For Daily Weather Simulation. U.S. Department of Agriculture, Agricultural Research Service, ARS-75, 49 pp.

The microcomputer program CLIMATE.BAS provides easy access to rainfall probabilities or simulated daily weather data for a State or region. Parameters for individual stations located within the region can be accessed directly, or they can be estimated for points between stations. In addition to precipitation, daily maximum temperatures, minimum temperatures and radiation can be simulated using a weakly stationary generating process conditioned on the precipitation process which is described by a Markov-chain/mixed-exponential model. The seasonal variations of parameters are described by Fourier series, providing a very parsimonious model.

KEYWORDS: Climate, Markov chain, microcomputer, precipitation, probability, simulation, solar radiation, temperature, weather

Computer printouts are reproduced essentially as supplied by the authors.

Copies of this publication may be purchased from the National Technical Information Service, 5285 Port Royal Road, Springfield, VA 22161.

ARS has no additional copies for free distribution.

PREFACE

This manual is designed to provide the theoretical background for the microcomputer program CLIMATE.BAS and to present an example analysis and program written for the State of South Dakota. We recommend that the nontechnical reader start with the introduction and then proceed to the section entitled EXAMPLE, on page 34. If the reader has an IBM-compatible personal computer with BASICA or a Health-Zenith Z-100 with ZBASIC, he or she will be able to run the programs on the enclosed diskette and discover their capabilities. The programs and data files on the enclosed diskette are in ASCII format and can be printed out for detailed examination. Any program revisions will be described in the file README.DOC on the enclosed diskette.

ACKNOWLEDGEMENTS

The assistance of K. Cumming, T. Econopouly, A. Haile, and F. Lopez in various phases of the data analysis and computer programming is gratefully acknowledged. W. Lytle, state climatologist, South Dakota State University, Brookings, SD, provided the precipitation data. Gratefully acknowledged are S. Anderson and L. Yohnka for the preparation of this manuscript for publication.

CONTENTS

NOTATIONS

c_{ik}	Amplitude of the kth harmonic of a Fourier series describing the probability of a transition from a dry day (i = 0) to a dry day or a wet day (i = 1) to a dry day.
C_j	Amplitude of the first harmonic for Fourier series representation of the mean or coefficient of variation of t_{max}, t_{min}, or r. (j = 1, 2, 3).
$C_j(n)$	coefficient of variation of t_{max} (j = 1), t_{min} (j = 2), or r (j = 3) on day n.
E { }	indicates expected value (mean).
$F_m(s)$	cumulative distribution function of total precipitation, s, in m days.
inf	infimum (smallest).
M	the number of years.
$\mathbf{M}_0$	3 X 3 matrix of lag zero correlation coefficients between t_{max}, t_{min}, and r.
$\mathbf{M}_1$	3 X 3 matrix of lag 1 correlation coefficients between t_{max}, t_{min}, and r.
M_i	maximum number of harmonics in the Fourier series describing the probability of a transition from a dry (i = 0) or wet (i = 1) to a dry day.
n	the day number, 1, 2,...365.
N(m)	number of wet days in an m-day period.
$P_{00}(n)$	probability of a transition from a dry day on day n-1 to a dry day on day n.
$P_{01}(n)$	probability of a transition from a dry day on day n-1 to a wet day on day n.
$P_{10}(n)$	probability of transition from a wet day on day n-1 to a dry day on day n.
$P_{11}(n)$	probability of a transition from a wet day on day n-1 to a wet day on day n.
$\bar{P}_{i0}$	annual mean probability of a transition.
P_x	number of horizontal pixels for microcomputer monitor.

P_y	number of vertical pixels for microcomputer monitor.
r	daily solar radiation [Langleys].
r_1, r_2	normally distributed random variables with mean = 0 and standard deviation = 1.
R_a	aspect ratio of computer monitor.
$s(m)$	Total precipitation in m days [L].
$s_j(n)$	standard deviation of t_{max} (j = 1), t_{min} (j = 2), or radiation (j = 3) on day n.
T	daily precipitation threshold [L].
$t_j(n)$	daily value of t_{max}(j = 1), t_{min}(j = 2), or radiation (j = 3).
t_{max}	daily maximum temperature [°F].
t_{min}	daily minimum temperature [°F].
u	horizontal coordinate.
u_1, u_2	uniformly distributed random variables on the interval [0, 1].
$u_j(n)$	Fourier series representation of t_{max}, t_{min}, or r on day n. (j = 1, 2, 3).
$\bar{u}_j$	Mean or coefficient of variation of the Fourier series for t_{max}, t_{min}, or r. (j = 1, 2, 3).
u_{max}	maximum horizontal coordinate on state map.
v	vertical coordinate.
v_{max}	maximum vertical coordinate on state map.
x	horizontal pixel coordinate for microcomputer monitor.
$X_\tau(n)$	random variable taking on the value 1 if day n of year τ is wet and 0 if it is dry.
y	observed precipitation.
y'	observed precipitation minus the threshold, T [L].
$Y(n)$	precipitation depth on day n [L].

$\alpha(n)$ weighting factor in the mixed exponential distribution for day n.

$\beta(n)$ parameter for mixed exponential distribution for day n. [L].

$\delta(n)$ parameter for mixed exponential distribution for day n. [L].

$\epsilon_j(n)$ normally distributed error term for t_{max} (j = 1), t_{min} (j = 2), or radiation (j = 3) for day n.

θ_{ik} phase angle in radians for the kth harmonic of a Fourier series describing the probability of a transition from a dry day (i = 0) or a wet day (i = 1) to a dry day.

λ mean of exponential distribution.

$\mu(n)$ mean of the precipitation greater than threshold T for day n [L].

$\mu_j(n)$ mean of t_{max} (j = 1), t_{min} (j = 2), or radiation (j = 3) on day n.

$\chi_j(n)$ vector of normalized residuals $[t_j(n) - \mu_j(n)]/s_j(n)$.

$\psi_0(m,k)$ probability that there are k wet days in an m-day period given that the prior day was dry.

$\psi_1(m,k)$ probability that there are k wet days in an m-day period given that the prior day was wet.

245

MICROCOMPUTER PROGRAM FOR DAILY WEATHER SIMULATION

D.A. Woolhiser, C.L. Hanson and C.W. Richardson

INTRODUCTION

Virtually every agricultural operation and many engineering activities are weather or climate dependent. For example, weather and climate information is essential for design of irrigation systems (Palmer et al. 1982), design of agricultural or urban drainage systems, selection of farm or construction machinery (Von Bargen 1967), and design of earthen covers for landfills and storage sites for low-level radioactive wastes (Lane 1984). In the short term, decisions must be made regarding what crop to plant, when to irrigate, when to apply fertilizer or pesticides, how much labor is required to harvest a crop, and many others. These decisions depend on past weather because the present stage of a crop (for example) is determined by the weather sequence since planting and they also depend on future weather. Unfortunately, we cannot predict future weather with certainty. Therefore, a certain amount of risk is involved in every decision made.

The concept that improved knowledge of risks due to weather would lead to improved decisions was a major reason for establishing the network of official and cooperative weather observation stations in the United States now operated by the National Weather Service (NWS) of the Department of Commerce and the networks of automatic weather recording stations now operated by many States.

The extensive weather data gathered by the NWS are available on magnetic tapes (now some are available on diskettes for microcomputer use). They are, however, somewhat cumbersome to use. The information content of the data can be summarized by statistical analyses and the results presented in tabular or graphical form. For example, results of extensive analyses performed by the regional climate committees of the State agricultural experiment stations during 1955-1967 were presented in a series of reports (Shaw et al. 1960, Dethier and McGuire 1961, Feyerherm et al. 1965, Gifford et al. 1967). Recent developments in microcomputer technology and lower costs of microcomputers combine to make new systems for the delivery of climatic information feasible.

In this report we describe the microcomputer program, CLIMATE.BAS, which provides easy access to rainfall probabilities or simulated daily weather for a location within a State

---

Woolhiser is a hydraulic engineer with USDA-ARS, Aridland Watershed Management Research, 2000 East Allen Road, Tucson, AZ 85719; Hanson is an agricultural engineer with USDA-ARS, Watershed Management Research, Northwest Watershed Research Center, Boise, ID 83705; and Richardson is an agricultural engineer with USDA-ARS Grassland, Soil and Water Research Laboratory, P.O. Box 748, Temple, TX 76503.

or region. Daily precipitation is described by a first-order Markov chain with precipitation amounts distributed as a mixed exponential. In addition, daily maximum temperatures, daily minimum temperatures and solar radiation can be simulated using a weakly stationary generating process first described by Matalas (1967) and adapted to daily weather by Richardson (1981). The seasonal variations of parameters are described by Fourier series providing a very parsimonious model. Through the interactive microcomputer program a user can access the information for a single station or can estimate weather characteristics for points between stations through a simple interpolation procedure. The program was developed for the State of South Dakota to provide an example.

This program is designed to supplement, not replace, actual climate data and real time weather data. One advantage of the simulation approach described here is that it doesn't require a great deal of computer memory or data storage capacity and can be used on rather modest microcomputer systems. Another advantage is that solar radiation can be estimated for locations where it has not been measured. Finally, the interpolation procedures allow estimates of daily weather characteristics at points between weather stations.

The objective of this manual is to provide sufficient technical information so the user can understand the assumptions of the model, the procedures of data analyses, and the modifications required in the microcomputer program to adapt it to another State or region.

THEORETICAL DESCRIPTION OF THE PRECIPITATION PROCESS

Daily Precipitation Occurrence

The occurrence or nonoccurrence of precipitation on day n of year τ can be represented by the random variable $X_\tau(n)$; $\tau = 1, 2, \cdots M$; $n = 1, 2, \cdots 365$ (the extra day in a leap year is ignored),

where

$$X_\tau(n) = \begin{cases} 0 & \text{if day n is dry.} \\ 1 & \text{if day n is wet.} \end{cases} \tag{1}$$

The dependence between wet and dry occurrences on successive days is described by a seasonally varying first-order Markov chain with transition probabilities $p_{ij}(n)$ $i = 0, 1$; $j = 0,1$,

where

$$p_{ij}(n) = P\{X_\tau(n) = j | X_\tau(n-1) = i\} \text{ for } n > 1, \text{ and} \tag{2}$$

$$p_{ij}(1) = P\{X_\tau(1) = j | X_{\tau-1}(365) = i\}$$

With this in mind we will drop the subscript τ in subsequent developments. Because $p_{i1}(n) = 1-p_{i0}(n)$, only two parameters are required for each day. Seasonal variations are accounted for by expressing the transition probabilities as a Fourier series:

$$p_{i0}(n) = \bar{p}_{i0} + \sum_{k=1}^{m_i} c_{ik}\sin(2\pi nk/365 + \theta_{ik}); \quad n = 1,2,\cdots 365 \tag{3}$$

where $i = 0$ or 1; m_i = the maximum number of harmonics required to describe the seasonal variability of the transition probability, $\bar{p}_{i0}$ is the annual mean value of the parameter, c_{ik} is the amplitude, and θ_{ik} is the phase angle in radians for the kth harmonic. The means, amplitudes, and phase angles were estimated by numerical optimization of the log likelihood function, as described by Woolhiser and Pegram (1979) and Roldan and Woolhiser (1982). Fourier series representations of parameters in a first-order Markov chain for precipitation have been used (among others) by Feyerherm and Bark (1965) who used least squares techniques for parameter estimation and by Stern and Coe (1984) who formulated the estimation problem as a generalized linear model to obtain maximum likelihood estimators.

The unconditional probability of a wet day on day n can be approximated by the following expression:

$$P\{X(n) = 1\} \approx [1 - p_{00}(n)]/[1 + p_{10}(n) - p_{00}(n)] \tag{4}$$

Distribution of Daily Precipitation

Daily precipitation amounts above a threshold, T, are described by the mixed exponential distribution (Smith and Schreiber 1974):

$$f_n(y') = \frac{\alpha(n)\ \exp[-y'/\beta(n)]}{\beta(n)} + \frac{[1 - \alpha(n)]\ \exp[-y'/\delta(n)]}{\delta(n)} \qquad (5)$$

where $y' = y-T$, the daily precipitation amount minus a threshold, T, provided $y > T$; $\alpha(n)$ = a weighting parameter with values between 0 and 1; and $\beta(n)$ and $\delta(n)$ are,the means of the smaller and the larger exponential distributions, respectively. Let $\mu(n)$ be the mean of $y'(n)$. It can be described in terms of the other parameters by the relation:

$$\mu(n) = \alpha(n)\beta(n) + (1-\alpha(n))\delta(n) \qquad (6)$$

The seasonal variations of these parameters are also represented by Fourier series, and the means, amplitudes, and phase angles were estimated by numerical maximization of the log likelihood function as described in the appendix. Significant harmonics were determined by the Akaike information criterion (AIC) (Akaike 1974).

Expected Annual Precipitation

Precipitation occurrence process, X(n), is assumed to be independent of the distribution of amounts. (This seems to be a good assumption; see Stern and Coe 1984, Woolhiser and Roldan 1982.) Under this assumption the expected total precipitation for m days, E{S(m)}, can be written as the sum

$$E\{S(m)\} = \sum_{n=1}^{m} E\{\ X(n)\ Y(n)\} = \sum_{n=1}^{m} E\{Y(n)\}E\{\ X(n)\} \qquad (7)$$

where Y(n) is the precipitation depth on day n.

Now

$$E\{X(n)\} = P\{X(n) = 1\} = P\{X(n-1) = 0\} \qquad (8)$$

$$p_{01}(n) + P\{X(n-1) = 1\}\ p_{11}(n)$$

and

$$E\{Y(n)\} = \alpha(n)\beta(n) + [1-\alpha(n)]\delta(n) + T \qquad (9)$$

Thus, for annual precipitation,

$$E\{S(m)\} = \sum_{n=1}^{365} [P\{X(n-1) = 0\}p_{01}(n) + P\{X(n-1) = 1\}p_{11}(n)] \qquad (10)$$

$$\cdot\ [\alpha(n)\beta(n) + \{1-\alpha(n)\}\delta(n) + T]$$

The extra day in a leap year is ignored.

Distribution of Total Precipitation in m Days

The total precipitation in m days can be written as

$$S(m) = \sum_{n=1}^{m} X(n)\ Y(n) \tag{11}$$

Thus, the distribution function of S(m) can be written as

$$F_m(s) = P\{S(m) \le s\} = P\{S(m) = 0\} + \sum_{j=1}^{m} P\{S(m) \le s | N(m) \tag{12}$$

$$= j\}P\{N(m)=j\}$$

where N(m) is the number of wet days in the m-day period.

An analytical expression for this distribution was derived by Todorovic and Woolhiser (1975) using results from Gabriel (1959) and Gabriel and Neuman (1962) for the Markov chain counting process and an exponential distribution for the daily precipitation:

$$F_m(s) = (1 - q_0 - Rd)(1 - q_0)^{m-1} \tag{13}$$

$$+ \sum_{k=1}^{m} [R\psi_1(m,k) + (1 - R)\psi_0(m,k)]\ \frac{\lambda^k}{k} \int_o^{s-kT} u^{k-1}\ e^{-\lambda u} du$$

where

(a) $\lambda = 1/\mu(n)$
where n is the day at or adjacent to the midpoint of period m.

(b) $q_0 = P_{01} = 1-P_{00}$

(c) $q_1 = P_{11} = 1-P_{10}$

(d) $d = q_1-q_0 = P_{00}-P_{10}$

(e) $R = p\ \{X_0=1\}$
where X_0 refers to the occurrence process on the day before the m day period.

$$\psi_1(m,k) = P\{N(m) = k|X_0 = 1\} \tag{14}$$

$$= q_1^k(1-q_0)^{m-k} + \sum_{c=1}^{C_1} \begin{pmatrix} k \\ a \end{pmatrix} \begin{pmatrix} m-k-1 \\ b-1 \end{pmatrix} \left(\frac{1-q_1}{1-q_0}\right)^b \left(\frac{q_0}{q_1}\right)^a$$

where

$$C_1 = \begin{cases} m + 1/2 - |2k - m + 1/2|, & \text{if } k < m \\ 0, & \text{if } k = m \end{cases}$$

$$a = \inf \{\nu \ ; \ \nu \geq 1/2(c-1)\}$$

$$b = \inf \{\nu \ ; \ \nu \geq 1/2 \ c\}$$

and

$$\psi_0(m,k) = P\{N(m) = k | X_0 = 0\} \tag{15}$$

$$= q_1(1 - q_0)^{m-k}$$

$$+ \sum_{c=1}^{C_0} \binom{k-1}{b-1} \binom{m-k}{a} \left(\frac{1-q_1}{1-q_0}\right)^a \left(\frac{q_0}{q_1}\right)^b$$

where

$$C_0 = \begin{cases} m + 1/2 - |2k - m + 1/2|, & \text{if } k > 0 \\ 0, & \text{if } k = 0 \end{cases}$$

The procedure used in this program to describe the multivariate process of maximum temperature, t_{max}, minimum temperature, t_{min}, and solar radiation, r, has been described by Richardson (1981). It is based on the weakly stationary generating process used by Matalas (1967) for generating stream-flow at multiple sites. The basic equation is

$$t_j(n) = \chi_j(n)\ s_j(n) + \mu_j(n) \tag{16}$$

where $t_1(n)$ is the daily value of t_{max}(on day n), $t_2(n)$ is t_{min} on day n, $t_3(n)$ is the value of r on day n, $s_j(n)$ is the standard deviation, and $\mu_j(n)$ is the mean of t_j for day n. The values of $\mu_j(n)$ and $s_j(n)$ are conditioned on whether the day was dry or wet, as determined by the Markov chain occurrence model. $\chi_j(n)$ is a vector of residuals obtained from the equation

$$\chi_j(n) = \mathbf{A}\chi_j\ (n-1) + \mathbf{B}\epsilon_j(n) \tag{17}$$

where $\chi_n(j)$ is a vector whose elements are the standardized residuals of t_{max}, t_{min}, and r, **A** and **B** are 3 X 3 matrices with elements defined to maintain the appropriate serial and cross correlation coefficients and ϵ_j is a vector of independent, normally distributed random variables with mean 0 and standard deviation of 1. The **A** and **B** matrices are given by

$$\mathbf{A} = \mathbf{M}_1\ \mathbf{M}_0^{-1} \tag{18}$$

$$\mathbf{B}\ \mathbf{B}^T = \mathbf{M}_0 - \mathbf{M}_1\ \mathbf{M}_0^{-1}\ \mathbf{M}_1^T \tag{19}$$

where the superscripts -1 and T denote the inverse and transpose, respectively. $\mathbf{M}_0$ and $\mathbf{M}_1$ are defined as

$$\mathbf{M}_0 = \begin{bmatrix} 1 & \rho_0(1,2) & \rho_0(1,3) \\ \rho_0(1,2) & 1 & \rho_0(2,3) \\ \rho_0(1,3) & \rho_0(2,3) & 1 \end{bmatrix} \tag{20}$$

$$\mathbf{M}_1 = \begin{bmatrix} \rho_1(1) & \rho_1(1,2) & \rho_1(1,3) \\ \rho_1(2,1) & \rho_1(2) & \rho_1(2,3) \\ \rho_1(3,1) & \rho_1(3,2) & \rho_1(3) \end{bmatrix} \tag{21}$$

where $\rho_0(j,k)$ is the correlation coefficient between variables j and k on the same day, $\rho_1(j,k)$ is the correlation coefficient between variable j and with variable k lagged 1 day, and $\rho_1(j)$ is the lag 1 serial correlation coefficient for variable j.

Richardson (1982) found that the correlation coefficients in matrices $\mathbf{M}_0$ and $\mathbf{M}_1$ showed little spatial variability for 31 locations in the United States. The average correlation coefficients given by Richardson (1982) and Richardson and Wright (1984) are

$$\mathbf{M}_0 = \begin{bmatrix} 1.000 & 0.633 & 0.186 \\ 0.633 & 1.000 & -0.193 \\ 0.186 & -0.193 & 1.000 \end{bmatrix} \tag{22}$$

$$\mathbf{M}_1 = \begin{bmatrix} 0.621 & 0.445 & -0.087 \\ 0.563 & 0.674 & -0.100 \\ 0.015 & -0.091 & 0.251 \end{bmatrix} \tag{23}$$

The **A** and **B** matrices can be computed for these values from equations (18) and (19).

$$\mathbf{A} = \begin{bmatrix} 0.567 & 0.086 & -0.002 \\ 0.253 & 0.504 & -0.050 \\ -0.006 & -0.039 & 0.244 \end{bmatrix} \tag{24}$$

$$\mathbf{B} = \begin{bmatrix} 0.782 & 0 & 0 \\ 0.328 & 0.637 & 0 \\ 0.238 & -0.341 & 0.873 \end{bmatrix} \tag{25}$$

Equation (16) can be written in the form

$$t_j(n) = \mu_j(n)[\chi_j(n)\ c_j(n) + 1] \tag{26}$$

where $c_j(n)$ is the coefficient of variation.

The seasonal changes in the means and coefficients of variation are represented by an equation of the form:

$$u_j(n) = \bar{u}_j + c_j \cos\left(0.0172(n - D_j)\right), \quad n = 1, 2, \cdots 365 \tag{27}$$

where u_j (n) is the value of the mean or coefficient of variation on day n, $\bar{u}_j$ is the annual mean, c_j is the amplitude of the first harmonic, and D is the phase angle in days. These variables were determined from 20 years of daily precipitation data for 31 U.S. locations, and are presented in maps and tables in Richardson and Wright (1984).

A flowchart for the program CLIMATE.BAS is shown in figure 1, and the program in BASIC is on the diskette in files CLIMATPC.BAS and CLIMATZ.BAS. In this section, the techniques used in the program will be explained in sufficient detail so that the user can modify it for another State or region.

Graphical Input for Maps

First obtain a map of the State or region at a convenient scale. Establish u and v coordinates, with the origin at the upper left corner and with v positive downward as shown in figure 2. Identify the points on the State boundary that are required to portray the boundary by using straight line segments. Using any convenient scale, find the u and v coordinates of these points and write them on the working copy of the map. Now determine the maximum values of u and v that you wish to display on the screen. These values, of course, will depend on the shape of the State and the amount of space required for labels, titles, scale of miles and white space. Let the maximum of u be u_{max} and the maximum of v be v_{max}. Establish the transformations u to x and v to y (where x and y are pixel numbers in horizontal and vertical directions) so that the largest dimension on the map coordinate system will fit on the screen. If $v_{max}/u_{max} \le (P_y-1)/[R_a(P_x-1)]$, the appropriate transformations are

$$x = (P_x-1)\ u/u_{max} \tag{28}$$

$$y = (P_x-1)\ R_a\ v/_{vmax} \tag{29}$$

where R_a is the aspect ratio, that is, the ratio of vertical to horizontal pixels to create a square on the screen, and P_x is the number of pixels in the horizontal direction. If $v_{max}/u_{max} > (P_y-1)/[R_a(P_x-1)]$, the vertical distance controls and the transformations are

$$x = (P_y-1)u/(R_a\ v_{max}) \tag{30}$$

$$y = (P_y-1)v/v_{max} \tag{31}$$

where P_y is the number of pixels in the vertical direction. For the Heath/Zenith Z-100, R_a = 0.4843, P_x = 640, and P_y = 225.

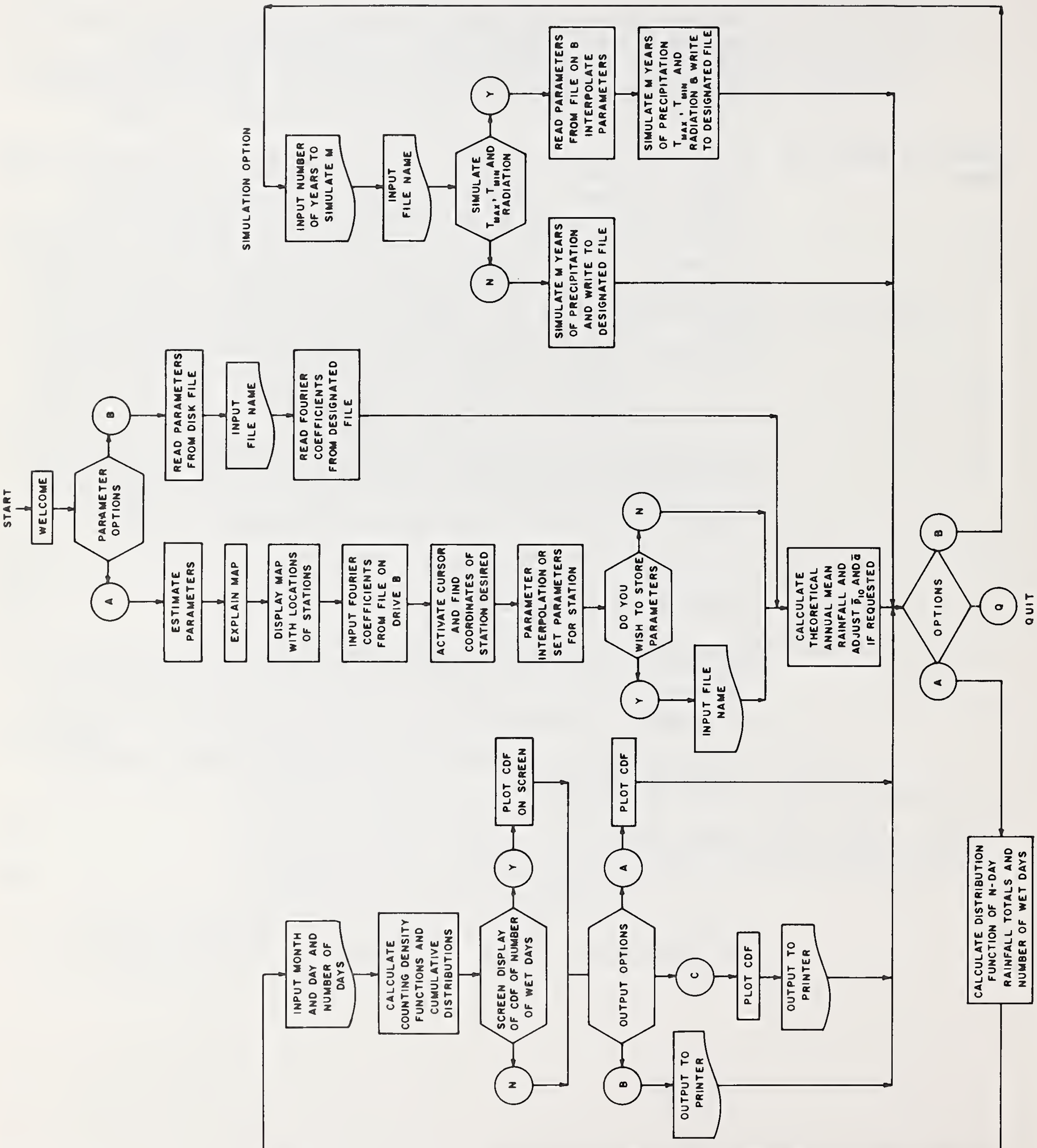


Figure 1.
Flowchart for program CLIMATE.BAS.

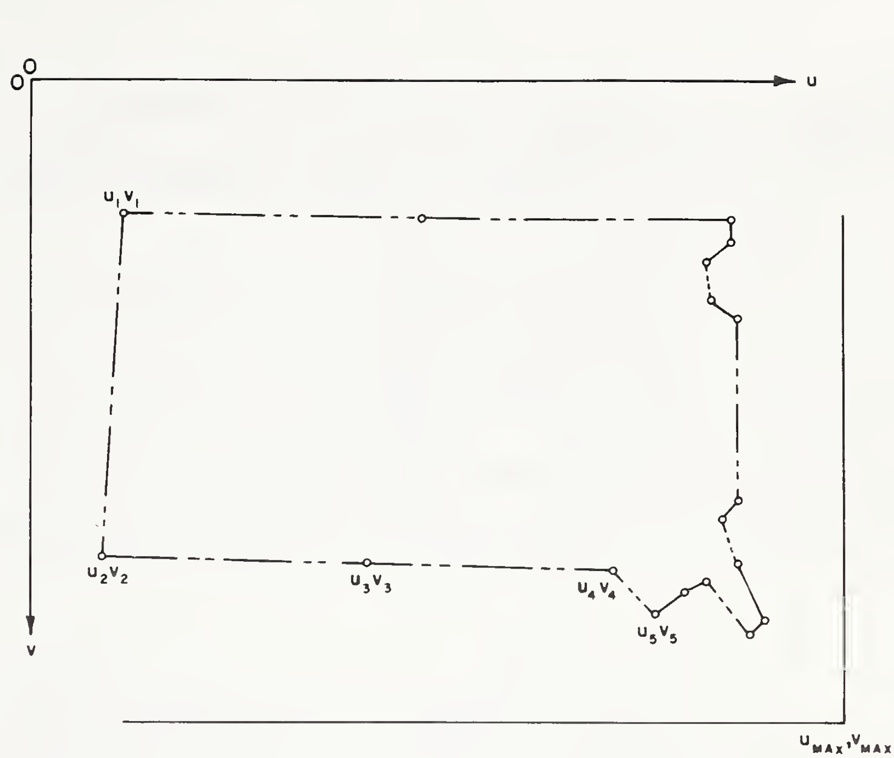


Figure 2.
Coordinate system for conversion to computer pixel coordinates.

The appropriate transformations should be made with a calculator, and the new coordinates written on the working copy of the map. The BASIC codes for drawing the map can then be written and tested.

Other modifications required in the BASIC code are--

1. Statement numbers 960, 1200, and 4560: Set the terminating index equal to the number of stations.

2. Statement 1100: Modify RMAX and RMIN such that they equal 100 and 30 miles, respectively. With the scale used for this program, 100 miles on the map was equal to 88.3 units. If the stations are more closely spaced than those for South Dakota, these radii can be reduced.

3. Statements 1470 through 1520 were included to account for the special case when the mountain station, Lead, is within 100 miles. They may be removed or modified.

4. Statement 210: Dimensions should be increased for X, Y, Z, R, and N% if there are more than 20 stations.

5. Statement 4780 must be modified to obtain the proper latitude.

Parameter Interpolation

There are two options for interpolation of parameters in this program:

1. Arithmetic average of parameters within a radius of 100 miles.

2. Nearest neighbor.

Woolhiser and Roldan (1986) compared the following methods for interpolating parameters:

1. Nearest neighbor.

2. Arithmetic average.

3. Spline function fit through the six nearest points.

4. Linear interpolation.

5. Kriging.

They found that for data from South Dakota, the arithmetic average of coefficients for the six nearest stations was superior to the other methods. The nature of the variogram was such that Kriging reduced to the average, that is, for the distance classes available, it showed a pure nugget effect (that is, no spatial dependence). Precipitation is strongly affected by orographic factors, so parameter averaging should not be used if adjacent stations differ widely in elevation. There is some evidence (Osborn et al. 1987) that the parameters in the Markov-chain/mixed exponential model can be adjusted to account for elevation, but more research is required before it can be incorporated into the microcomputer program.

In this program, when the coordinates of the cursor are calculated, the distance to each station is determined, and the stations within 100 miles are identified. Circles with radii of 30 and 100 miles are projected on the screen as an aid to the user's judgment. If the nearest station is closer than 30 miles, the user is asked if he wishes to use the parameters for the nearest station. If the response is yes, the method becomes a nearest neighbor estimate. If the answer is no, the estimated parameters will be averages of those for the stations within the 100-mile radius. The user has the option to omit any of these stations. If the Black Hills station, Lead, is within the 100-mile radius, a warning appears, because it is at a much higher elevation than the other stations, and has much more precipitation and cooler temperatures. If the user was estimating parameters for a station in the plains near Rapid City, he would probably omit Lead. If, on the other hand, he was estimating parameters for a station in the Black Hills, between Lead and Rapid City, he might eliminate all the plains stations except Rapid City. In the averaging procedure, all amplitudes, including zero, are averaged, but phase angles for nonsignificant harmonics are not included in the average.

Parameter Identification

The parameter matrix Z(I,J,K), read in statement 1020 from the sequential file "B:DAKWEA.DAT", consists of the means, amplitudes, and phase angles of six harmonics for the Fourier series representation of p_{00}, p_{10}, β, and μ in equations (2), (3), (5), and (6). The index I refers to the station number, J refers to the parameter, and K refers to the mean, amplitude or phase angle, depending on its value (see table 1).

These parameters were identified from 40 years of daily precipitation data using the program AGUA46, which provides approximate maximum likelihood estimates of the means, amplitudes, and phase angles. Procedures were developed by Woolhiser and Pegram (1979), Roldan and Woolhiser (1982), and Woolhiser and Roldan (1986) and are described in the appendix. Significant harmonics were determined by the Akaike information criterion.

Table 1.
Precipitation parameter indices

J	Parameter
1	p_{00}
2	p_{10}
3	β
4	μ

K	Fourier Parameter
1	Annual mean
2	Amplitude of first harmonic
3	Phase angle of first harmonic
4	Amplitude of second harmonic
5	Phase angle of second harmonic
6	Amplitude of third harmonic
7	Phase angle of third harmonic
8	Amplitude of fourth harmonic
9	Phase angle of fourth harmonic
10	Amplitude of fifth harmonic
11	Phase angle of fifth harmonic
12	Amplitude of sixth harmonic
13	Phase angle of sixth harmonic

Parameter Estimation for Temperature and Radiation

The parameters W(1) through W(12) read in statement 4570 from the sequential file "B:DAKW.ASC" are required to describe daily mean values of t_{max}, t_{min}, and r. These parameters are defined in table 2.

These parameters are read for each station in the same order as the precipitation stations and are averaged using the same procedures described for precipitation.

The parameter values in the file DAKW.ASC were estimated by linear interpolation on the maps provided by Richardson and Wright (1984), figures A1-A12[1]. The temperature parameters for Lead are unreliable because they have not been adjusted for elevation. Users should be aware of the potential bias in simulated temperatures if there is considerable topographic relief. Richardson and Wright (1984) describe a procedure to correct simulated temperatures by adjusting them to maintain monthly mean temperatures obtained from other sources. A similar procedure should be added to CLIMATE.BAS for regions where linear interpolation on the maps provided by Richardson and Wright (1984) is inappropriate due to topographic effects.

Distribution Function for m-Day Rainfall

The program provides an analytic solution for the distribution function $F_m(s)$, equation (13), using solutions for equations (14) and (15) for $\psi_1(k,m)$ and $\psi_0(k,m)$ for $1 \le m < 30$. The parameters are all treated as constants within the m-day period, and are estimated for the midpoint of the period. The parameter λ, for the exponential distribution, is estimated from the mean of the mixed exponential, that is,

$$\lambda = \alpha\beta + (1 - \alpha)\delta. \tag{32}$$

The integral in equation (13)

$$\int_o^{s-kT} u^{k-1}\, e^{-\lambda u}\, du \tag{33}$$

is solved recursively using the relationship found in standard tables of integrals:

$$\int x^m \exp(ax)dx = (x^m \exp(ax)/a) - (m/a) \int x^{m-1} \exp(ax)\, dx \tag{34}$$

Table 2.
Temperature and radiation parameter definitions

Symbol	Definition
W(1)	Annual mean of t_{max} for dry days (°F)
W(2)	Amplitude of first harmonic for t_{max}, wet and dry days (°F)
W(3)	Annual mean of the coefficient of variation of t_{max}, wet and dry days.
W(4)	Amplitude of first harmonic of coefficient of variation of t_{max}, wet and dry days.
W(5)	Annual mean of t_{max} for wet days (°F)
W(6)	Annual mean of t_{min} for wet and dry days (°F)
W(7)	Amplitude of first harmonic of t_{min} for wet and dry days (°F)
W(8)	Annual mean of coefficient of variation of t_{min}, wet and dry days (°F)
W(9)	Amplitude of first harmonic of coefficient of variation of t_{min}, wet and dry days (°F)
W(10)	Annual mean of daily solar radiation on dry days (Langleys)
W(11)	Amplitude of first harmonic of daily solar radiation, wet and dry days (Langleys)
W(12)	Annual mean of daily solar radiation on wet days (Langleys)

Parameter Adjustment to Correct Mean Annual Precipitation

When the parameters for a station have been estimated by averaging those of surrounding stations, the theoretical annual average precipitation as calculated by equation (10) may be slightly different from the estimated annual precipitation obtained by interpolation on an isohyetal map. An option within the program allows the parameters $\bar{\alpha}$ and $\bar{p}_{10}$ to be adjusted by a Newton-Raphson iterative procedure so that the theoretical mean is within ± 0.1% of the known average annual precipitation. The corrections are partitioned equally between α and p_{10} according to the equations

$$0.5F + \frac{\partial F}{\partial \bar{p}_{10}} \Delta \bar{p}_{10} = 0 \tag{35}$$

$$0.5F + \frac{\partial F}{\partial \alpha} \Delta \bar{\alpha} = 0 \tag{36}$$

where F is the difference between the theoretical and known average precipitation, and $\Delta \bar{p}_{10}$ and $\Delta\alpha$ are corrections. The derivatives are approximated by writing equation (10) with the parameters p_{00}, p_{10}, α, β, and δ assumed to be constants and taking the partial derivatives

$$\frac{\partial F}{\partial \bar{p}_{10}} = \frac{1-\bar{p}_{00}}{(1+\bar{p}_{10}-\bar{p}_{00})^2} \left[\bar{\alpha}\,\bar{\beta} + (1-\bar{\alpha})\bar{\delta} \right] 365 \tag{37}$$

$$\frac{\partial F}{\partial \bar{\alpha}} = - \left(\frac{1-\bar{p}_{00}}{1+\bar{p}_{10}-\bar{p}_{00}} \right) \left(\bar{\beta} - \bar{\delta} \right) 365 \qquad (38)$$

where the bar above the parameter indicates the constant value in the Fourier Series expression. $\bar{\alpha}$ and $\bar{p}_{10}$ were chosen for adjustment because they typically have greater variances than the other parameters.

Simulation Procedures

Markov Chain Process

Preceding day dry: A uniform random number, u, $(0 \le u \le 1)$ is generated. If $u < p_{00}(n)$, day n is dry; if $u \ge p_{00}(n)$, day n is wet.

Preceding day wet: A uniform random number, u, is generated. If $u < p_{10}(n)$, day n is dry; if $u \ge p_{10}(n)$, day n is wet.

Mixed Exponential Distribution

If day n is wet, a new uniform random number is generated. If $u < \alpha(n)$, the depth is generated from an exponential distribution with mean $\beta(n)$ using the transformation

$$y = -\beta(n) \log u + 0.008 \qquad (39)$$

If $u \ge \alpha(n)$, the depth is generated from an exponential distribution with mean $\delta(n)$:

$$y = -\delta(n) \log u + 0.008 \qquad (40)$$

The depth 0.008 inch is added because the raw data are transformed by subtracting the threshold depth 0.008 so the mixed exponential has a lower bound of nearly zero.

T_{max}, T_{min}, and Radiation

The residuals in equation (17) are generated using the transformations (Naylor et al. 1966)

$$r_1 = [-2 \log u_1 \cos(2\pi u_2)]^{1/2} \qquad (41)$$

$$r_2 = [-2 \log u_1 \sin(2\pi u_2)]^{1/2}$$

where r_1 and r_2 are independent normal (0,1) random variables, and u_1 and u_2 are independent, uniformly distributed random variables. Because three normally distributed residuals are required for each day, six random variables are generated at one time, and are used for 2 days. In the current version of the program the random number generator in BASIC is used to generate u_1 and u_2.

System Requirements and Model Performance

This program has been run successfully on a Heath/Zenith Z-100 with two disk drives, a color monitor and 192 K ram under ZBASIC (program CLIMATZ.BAS) and on a COMPAQ, TELEX or IBM-PC with two disk drives, and a monochrome monitor under BASICA (program CLIMATPC.BAS). The operating systems were DOS version 2 or greater. Both programs read parameter files from disk B, (statements 950 and 4550). A comparison of parameter estimation and simulation run times is shown in table 3.

Table 3.
Comparison of run times for CLIMATE.BAS

	COMPAQ 386 16 MHZ	TELEX 1260 8 MHZ	Heath/Zenith Z-100 4.7 MHZ
Parameter estimation (precipitation)	14 sec	44 sec	120 sec
Simulation of precipitation only, written to diskette.	3.7 sec/yr	4.8 sec/yr	15 sec/yr
Simulation of precipitation, temperature, and radiation. Written to diskette.	94.5 sec/yr	115.5 sec/yr	330 sec/yr
Simulation of precipitation only. Written to hard disk.	1.4 sec/yr	4.3 sec/yr	-----
Simulation of precipitation, temperature, and radiation. Written to hard disk.	34.0 sec/yr	106.8 sec/yr	-----

VALIDATION OF THE DAILY PRECIPITATION MODEL

Validation implies an examination of the statistical characteristics of model output to determine if the model structure and parameter estimation techniques preserve those properties of the precipitation process deemed most important.

To demonstrate the degree of fit between simulated precipitation sequences and the historical data, we selected four stations representing the extremes of South Dakota climate. These stations are Aberdeen, Lead, Redig, and Yankton. Forty years of record were simulated for each station.

Annual Statistics

Historical and simulated annual mean precipitation and the standard deviations are shown in table 4 and histograms of the historical and simulated annual precipitation sequences are shown in figure 3. The model does very well in preserving the annual mean, but simulated standard deviations are biased low. Some of this bias may be due to the particular random number seed used in the simulations with the microcomputer. In a study of parameter identifiability, 10 simulations of 40-year records were performed for Aberdeen. The historical standard deviation was exceeded two times; and in all cases the null hypothesis, Ho: $\sigma_o^2 = \sigma_s^2$, could not be rejected using the F test at the 0.05 level, where σ_o^2 and σ_s^2 are the variances of the historical and simulated records, respectively. This tendency of daily precipitation models to underestimate the variance of annual total precipitation has been noted previously (Zucchini and Adamson 1984). They attributed this underestimation to "the model's inability to preserve the frequency of extreme n-day rainfalls in cases where these are associated with weather generating processes that are distinct from those that generate the bulk of the rainfall."

Table 4.
Annual precipitation statistics, 40-year record

	Historical		Simulated		
Station	Mean (in)	Std. dev. (in)	Mean (in)	Std. dev. (in)	F[1]
Aberdeen	19.97	4.20	19.87	3.36	1.56
Lead	24.20	5.74	24.19	3.59	2.55
			23.68	3.89	2.17
Redig	13.55	3.56	13.59	2.39	2.22
Yankton	22.82	6.49	22.64	3.99	2.64
			22.96	4.69	1.91

[1] $F_{0.05} = 1.69$

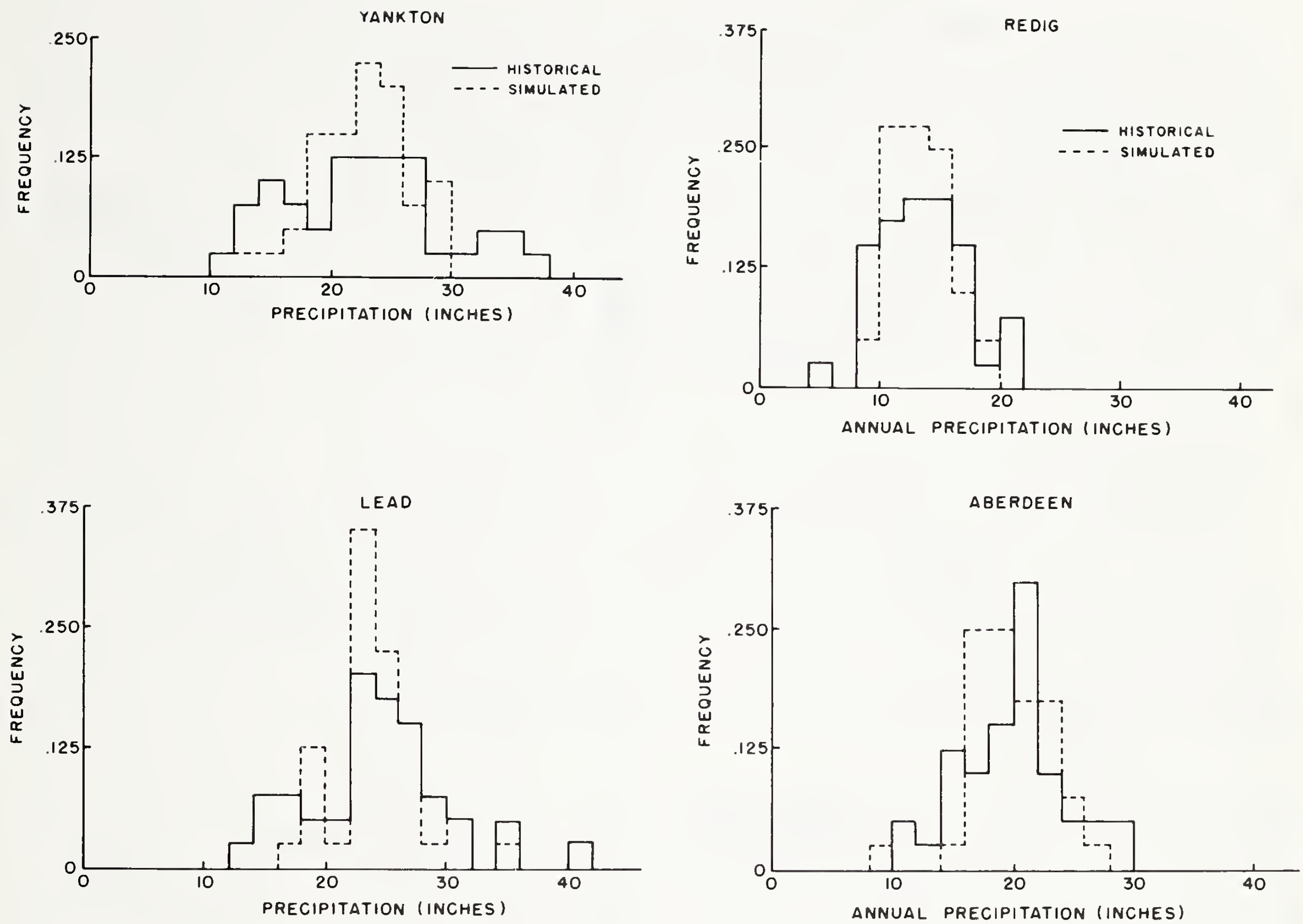


Figure 3.
Histograms of historical and simulated annual precipitation. Yankton, Lead, Redig, and Aberdeen, SD (m = 40 yr).

Statistics for 14-day Periods

The fidelity of the model in preserving the seasonal variability of the precipitation process is best demonstrated by examining some statistics for 14-day periods. The historical, simulated, and theoretical mean numbers of wet days for each of 26 periods are shown in figure 4. It is clear from this figure that the Markov chain model does an excellent job in preserving the mean number of wet days.

The historical and simulated mean and standard deviation of the precipitation per wet day are shown for each period in figures 5 and 6. The historical and simulated statistics compare very well, with no obvious tendency for either underestimating or overestimating the standard deviation.

The fidelity of the combination of the Markov chain and the mixed exponential model can be demonstrated by examining characteristics of total precipitation for n-day periods. The

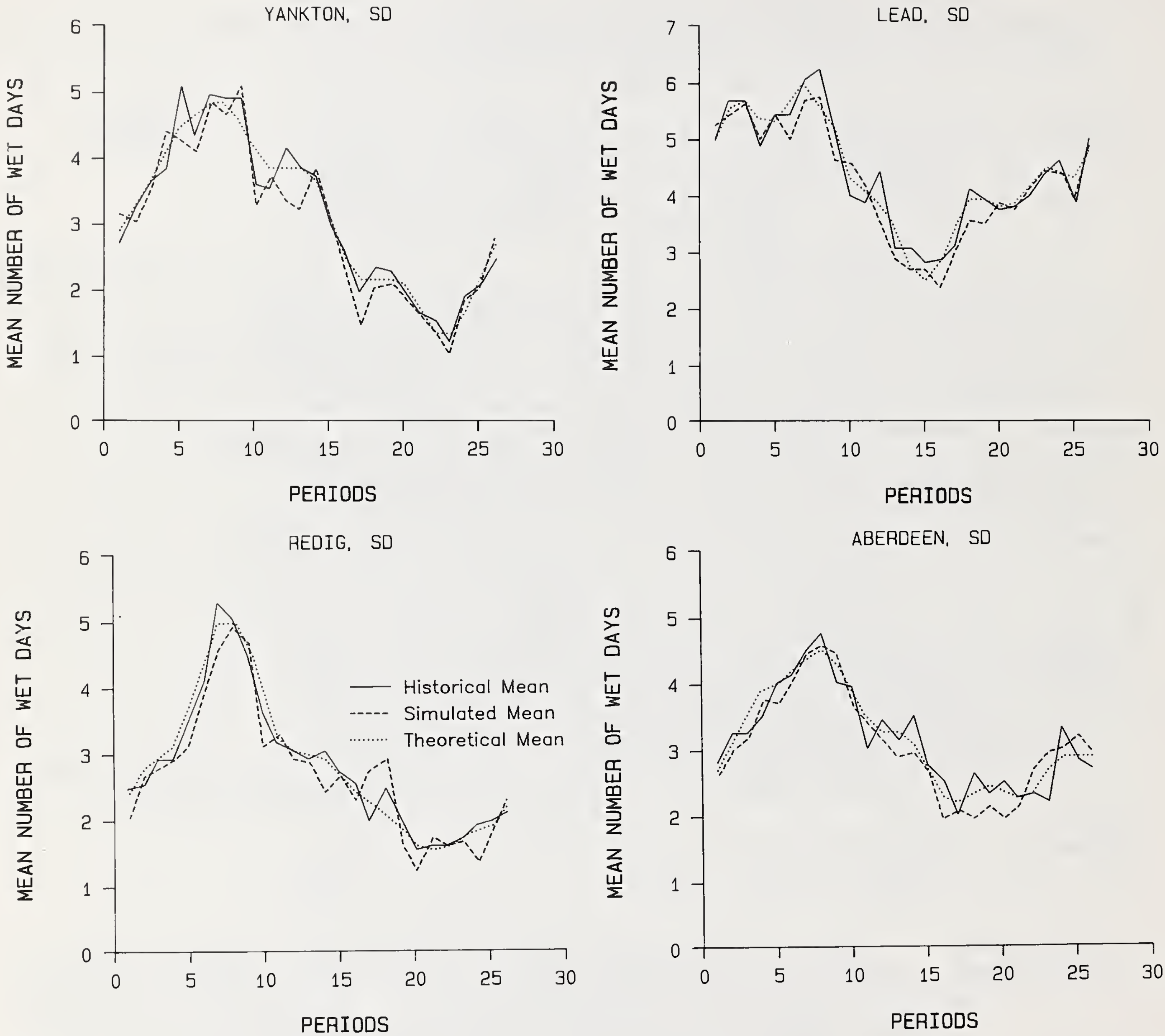


Figure 4.
Historical, simulated and theoretical mean number of wet days for 14-day periods.

historical and simulated means and standard deviations of 14-day precipitation are shown in figures 7 and 8. Again the fit is excellent. Finally the historical and simulated empirical distribution functions for three periods are compared in figures 9 and 10. These periods 1, (Mar. 1-14), 6 (May 10-23), and 14 (Aug. 30-Sept. 12) were chosen a priori. The Kolmogorov-Smirnov two-sample goodness of fit test was used to test the null hypothesis that the historical and simulated sample came from the same distribution. This hypothesis could not be rejected at the 0.05 level for all three periods for the four stations. The analytical approximation to the distribution function of 14-day precipitation (eq. 13) is also plotted in figures 9 and 10. It is clear that this approximation provides a very good fit to both historical and simulated data.

Some selected distribution functions of precipitation depths per wet day are shown in figures 11 and 12. The null hypothesis that the historical and simulated samples came from the same distribution was rejected for 2 out of 12 cases at the 0.05 level using the two sample K-S test. This is still acceptable, since under the null hypothesis, the probability of 2 or more rejections in 12 tests is 0.118. The rejections occurred for period 14 for Redig and period 6 for Aberdeen. An examination of figures 5 and 6 reveals that during these periods the mean and variance of the precipitation depth per wet day are changing rapidly, so this may also be a contributing factor.

From this brief validation study, it appears that one should use caution in the use of daily precipitation data generated with CLIMATE.BAS in studying annual phenomena (annual droughts for example). However the Markov-chain/mixed exponential model appears to do an excellent job of preserving important statistics within the year and can be recommended to provide input to models that have memories shorter than 1 year. The procedures for simulating maximum and minimum temperature and radiation were tested extensively by Richardson and Wright (1984) and were found to be satisfactory for the stations they analyzed. However, we did find considerable variation in the mean number of wet days per year for the South Dakota stations due to time of observation and other factors. It doesn't seriously affect precipitation, but it could bias simulated temperature and radiation if the station being simulated has many fewer wet days than the stations used by Richardson and Wright (1984) to obtain their temperature and radiation parameters.

Extremes of temperature in simulated data may also be unreliable because of microclimate factors as well as the simulation technique. Therefore we do not recommend using simulated data for estimating frost probabilities. Because of the strong dependence of both precipitation and temperature on elevation, interpolation procedures should not be used in mountainous terrain. For South Dakota this means that temperature and radiation simulated for Lead will be incorrect because the parameters were estimated for stations at lower elevations.

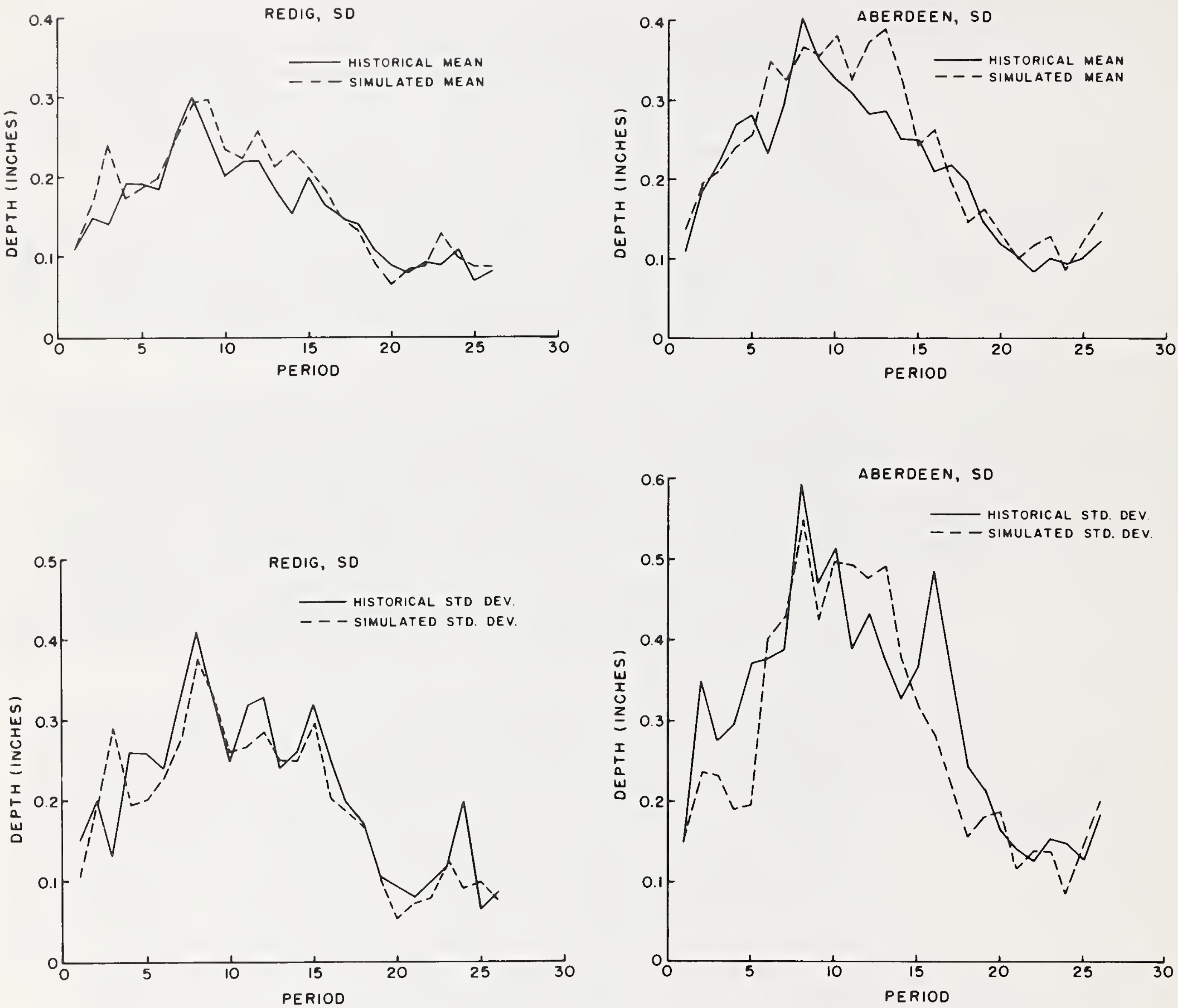


Figure 5.
Historical and simulated mean and standard deviations of precipitation per wet day for Redig and Aberdeen, SD.

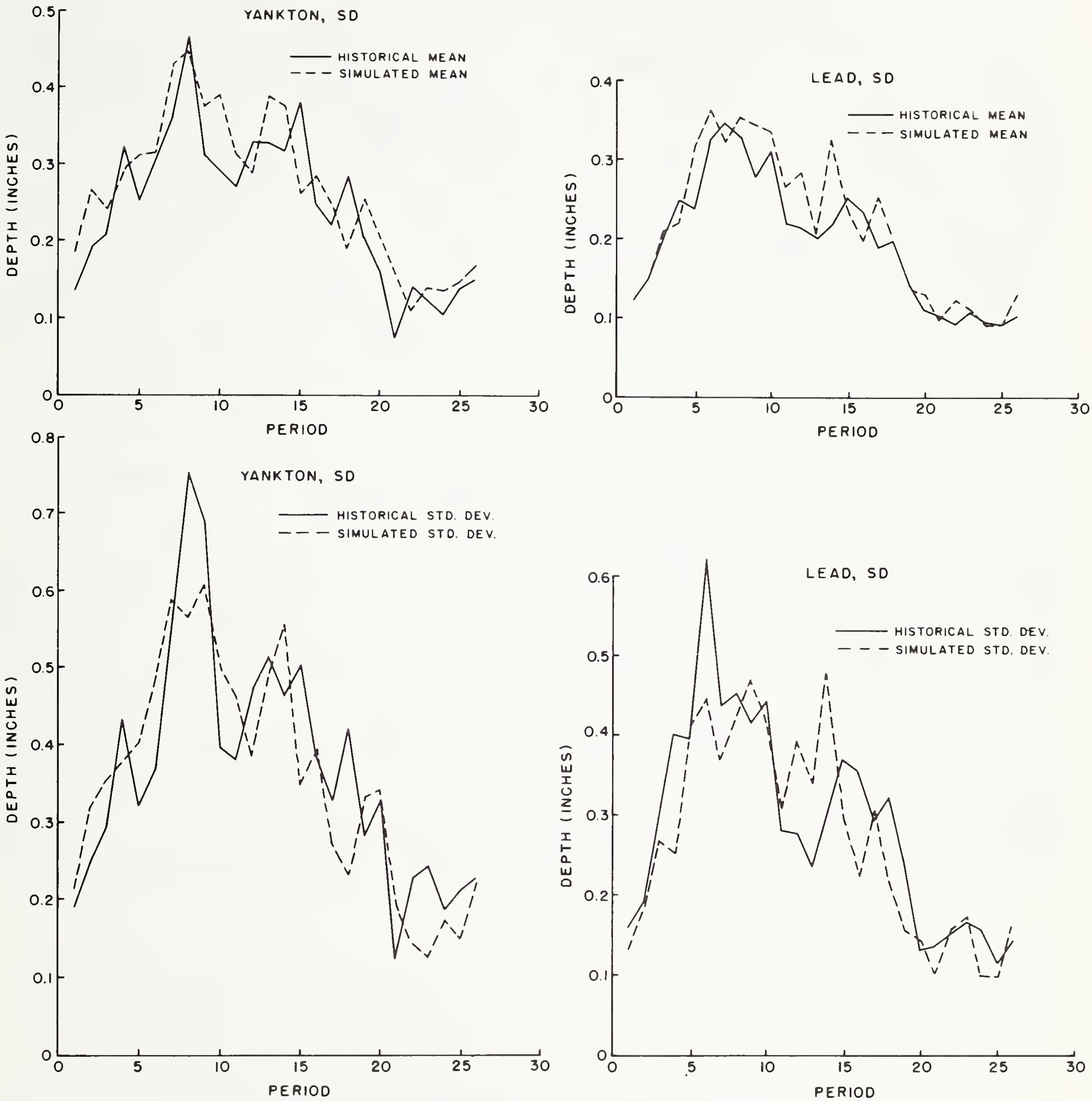


Figure 6.
Historical and simulated mean and standard deviations of precipitation per wet day for Yankton and Lead, SD.

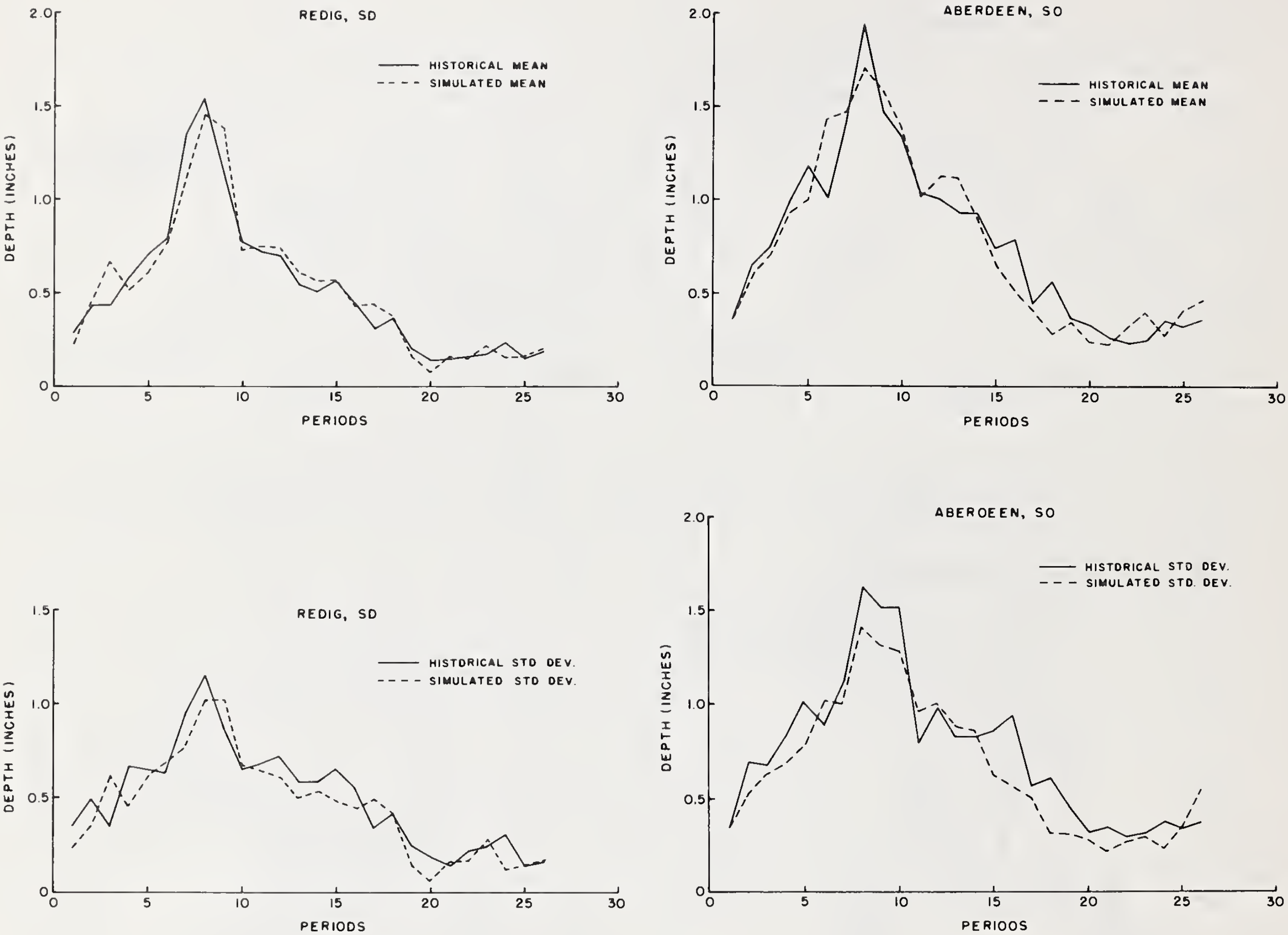


Figure 7.
Historical and simulated mean and standard deviation of total precipitation in 14-day periods for Redig and Aberdeen, SD.

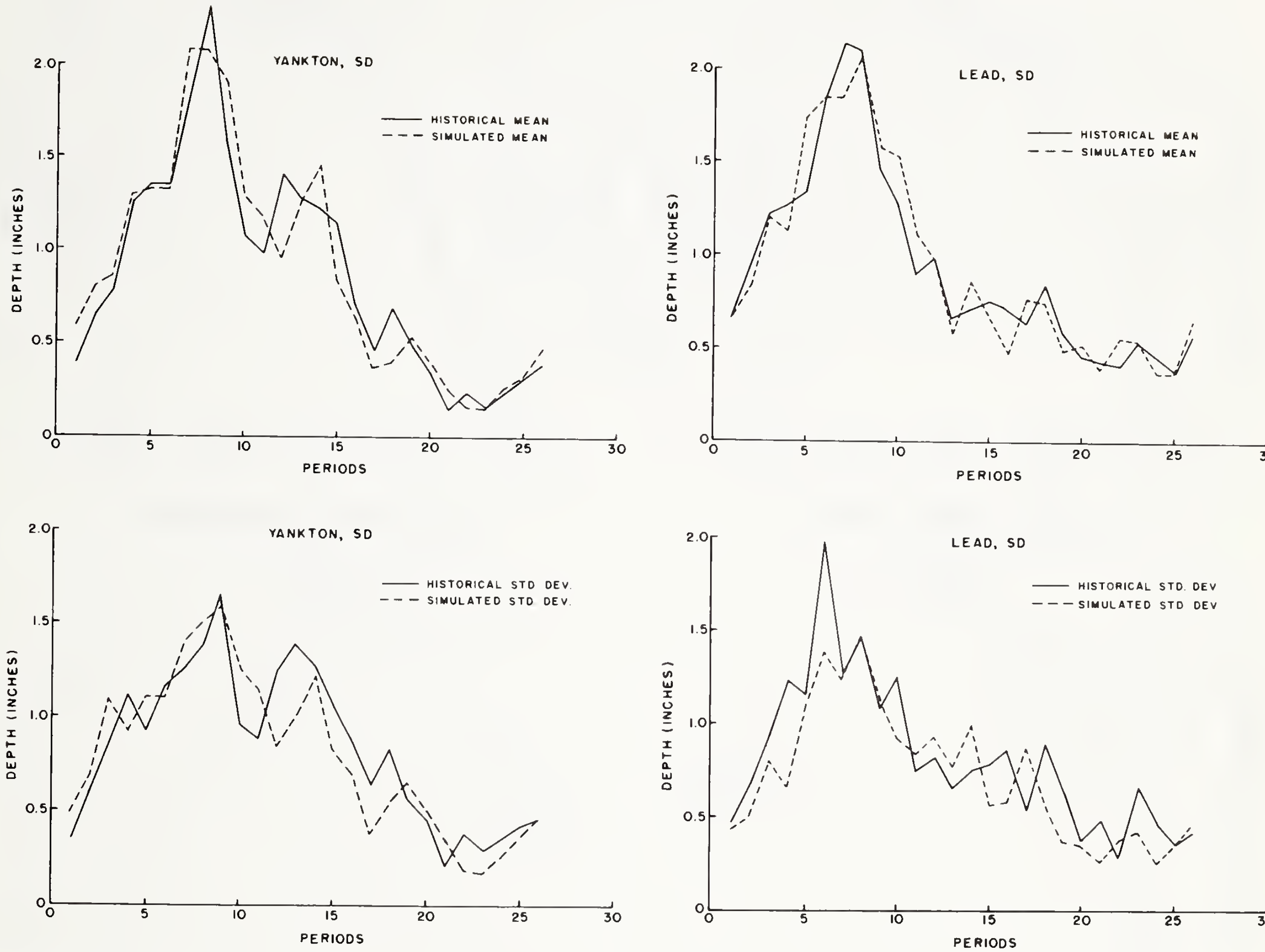


Figure 8.
Historical and simulated mean and standard deviation of total precipitation in 14-day period for Yankton and Lead, SD.

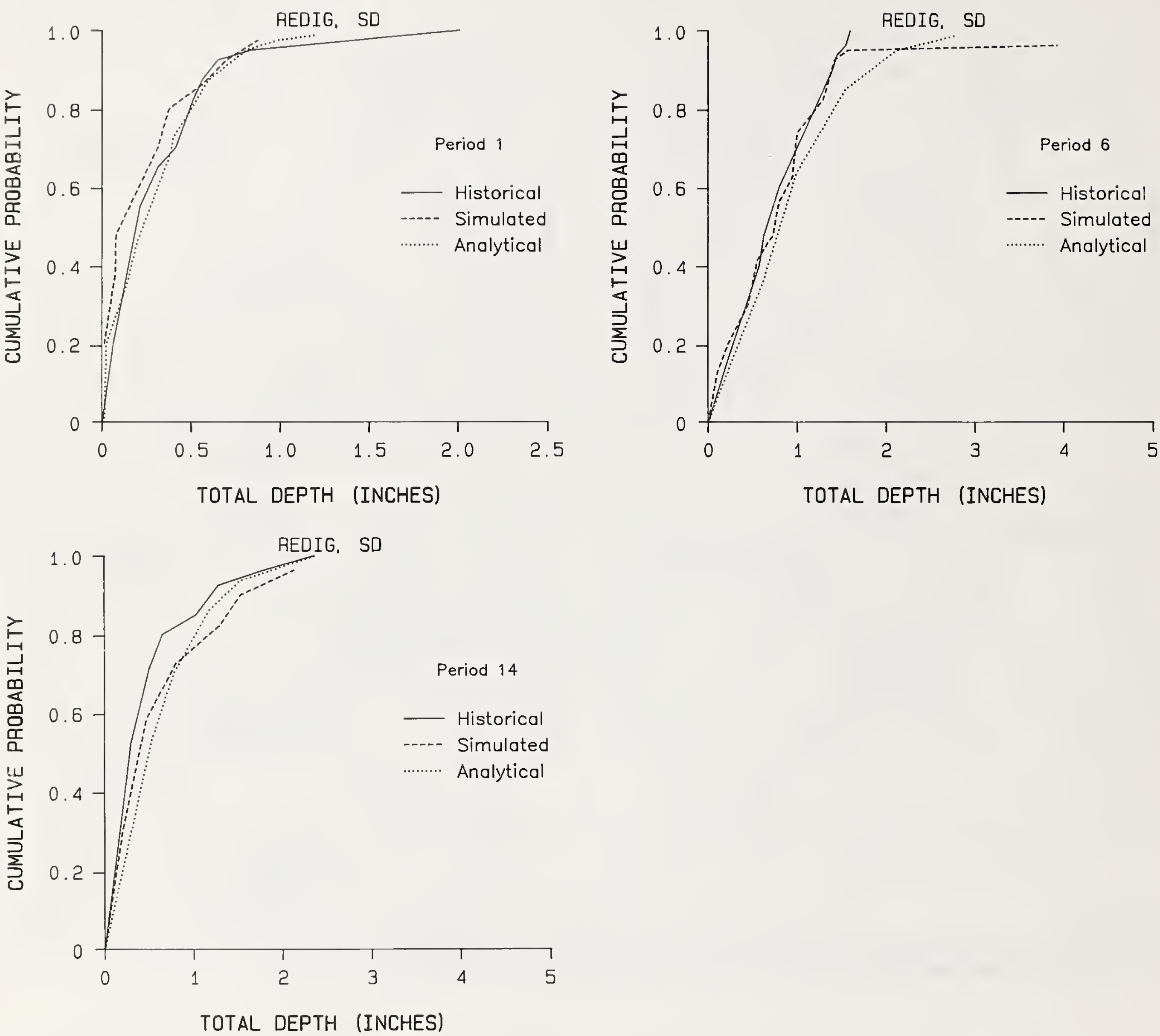
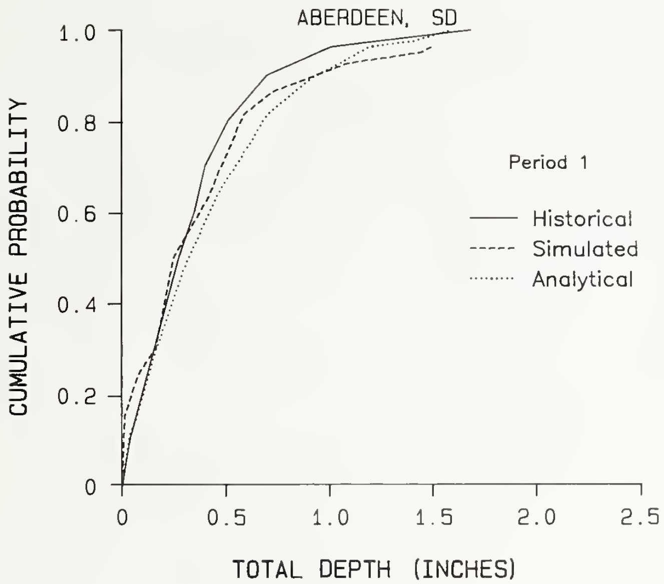
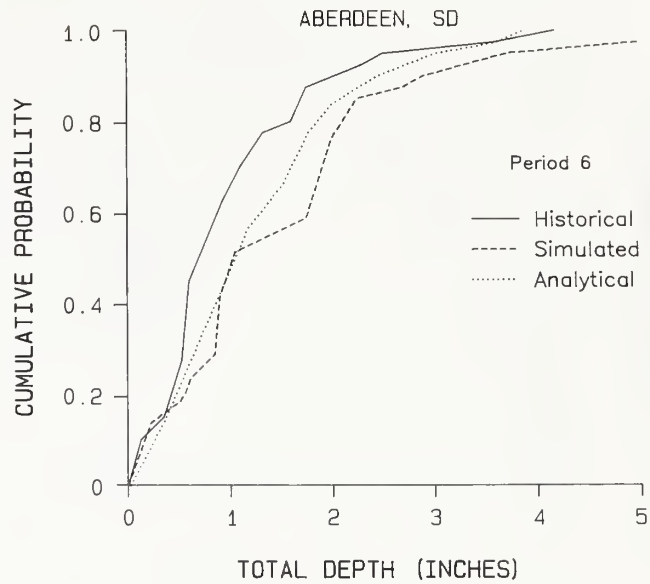


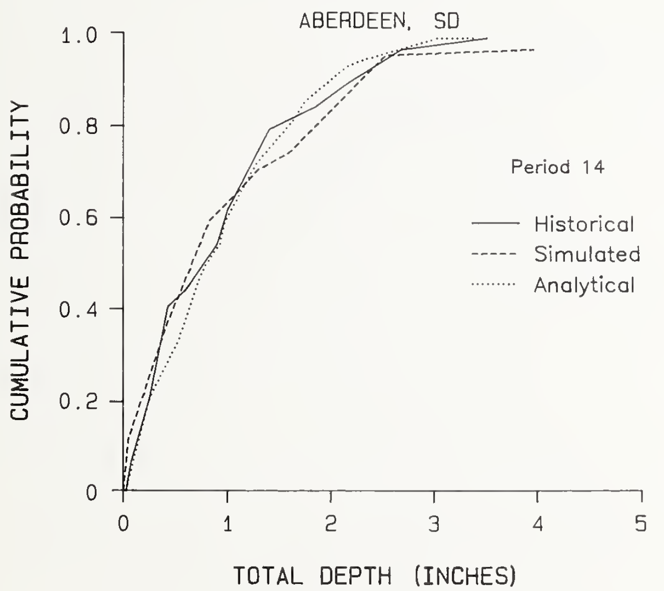
Figure 9.
Historical, simulated, and analytical cumulative distribution functions of total rainfall for three 14-day periods in Redig and Aberdeen, SD.



ABERDEEN, SD
CUMULATIVE PROBABILITY
TOTAL DEPTH (INCHES)
Period 1
Historical
Simulated
Analytical



ABERDEEN, SD
CUMULATIVE PROBABILITY
TOTAL DEPTH (INCHES)
Period 6
Historical
Simulated
Analytical



ABERDEEN, SD
CUMULATIVE PROBABILITY
TOTAL DEPTH (INCHES)
Period 14
Historical
Simulated
Analytical

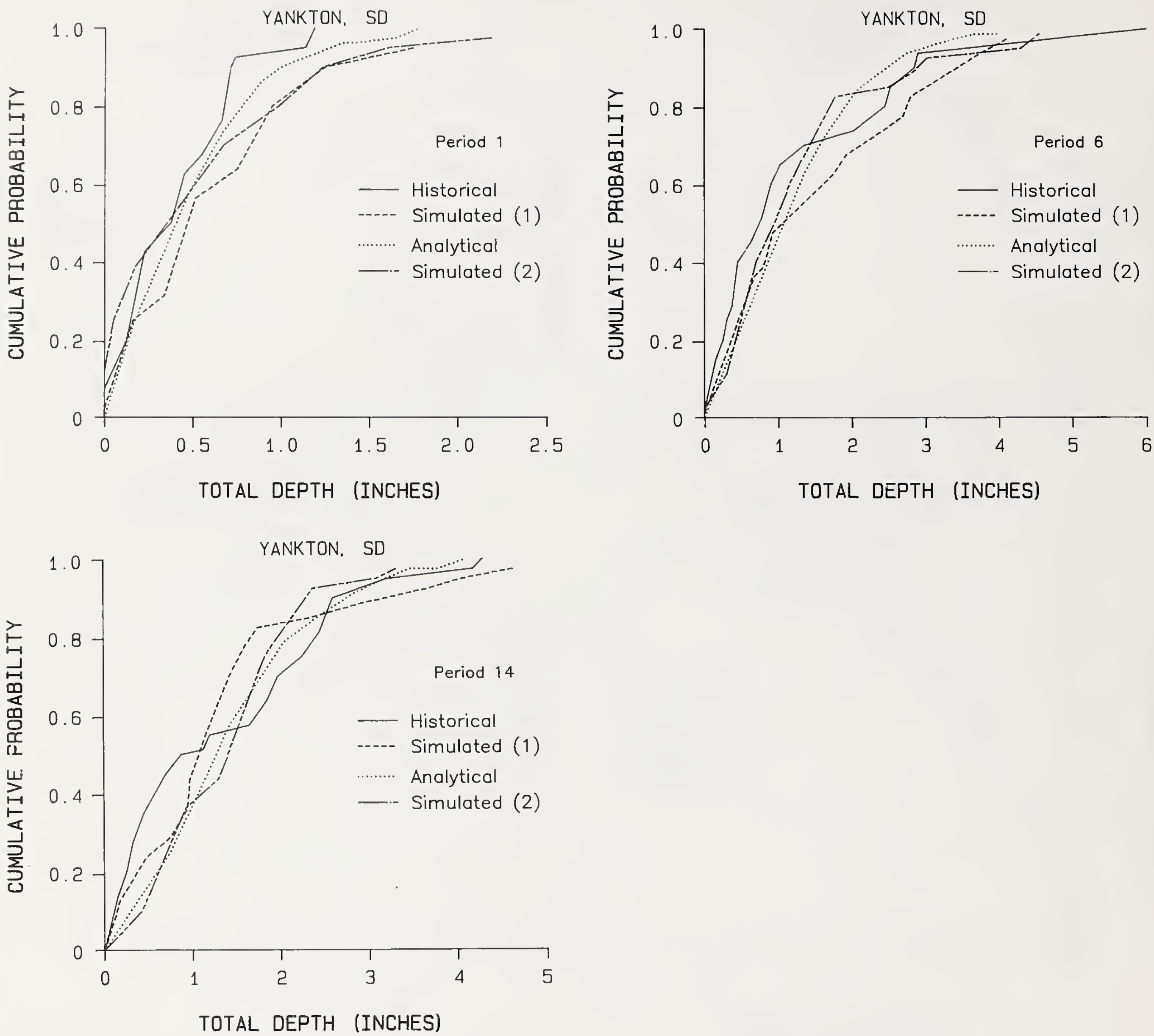
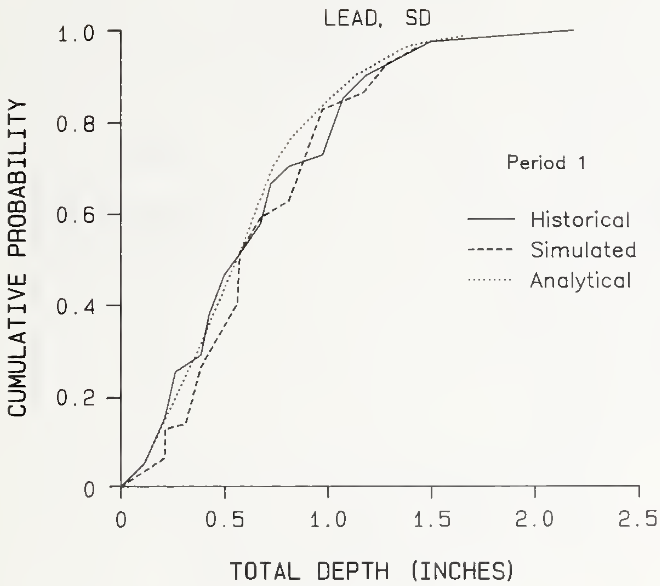
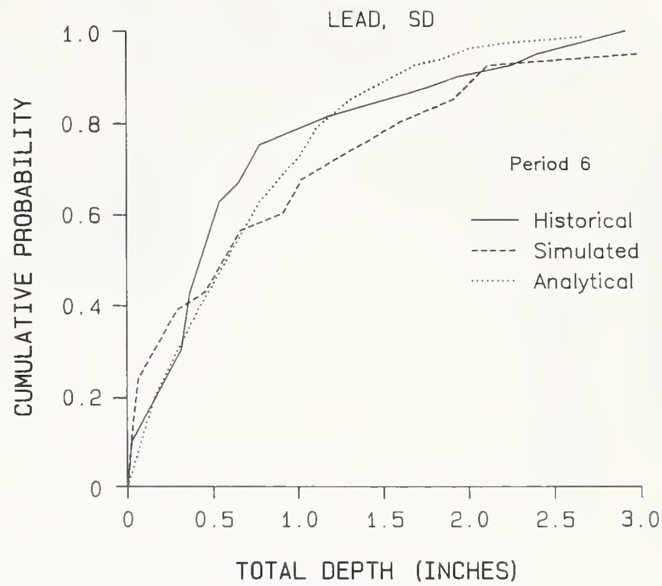
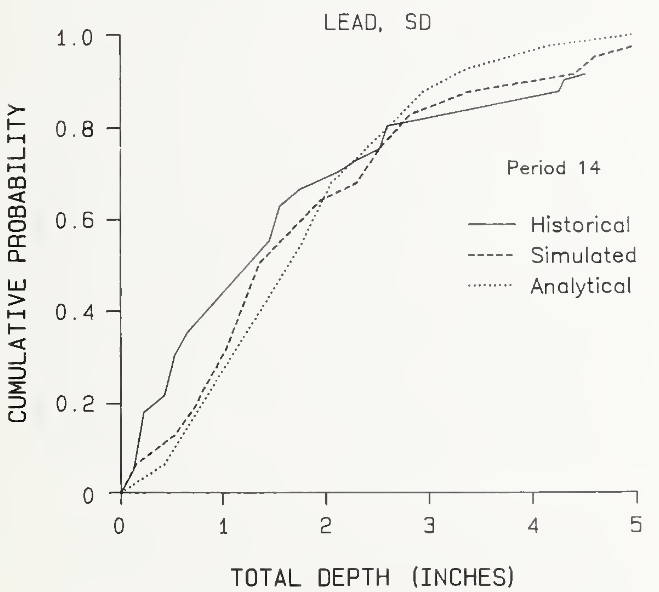


Figure 10.
Historical, simulated and analytical cumulative distribution functions of total rainfall for three 14-day periods in Yankton and Lead, SD.


LEAD, SD
CUMULATIVE PROBABILITY
1.0
0.8
0.6
0.4
0.2
0
0
0.5
1.0
1.5
2.0
2.5
TOTAL DEPTH (INCHES)
Period 1
Historical
Simulated
Analytical


LEAD, SD
CUMULATIVE PROBABILITY
1.0
0.8
0.6
0.4
0.2
0
0
0.5
1.0
1.5
2.0
2.5
3.0
TOTAL DEPTH (INCHES)
Period 6
Historical
Simulated
Analytical


LEAD, SD
CUMULATIVE PROBABILITY
1.0
0.8
0.6
0.4
0.2
0
0
1
2
3
4
5
TOTAL DEPTH (INCHES)
Period 14
Historical
Simulated
Analytical

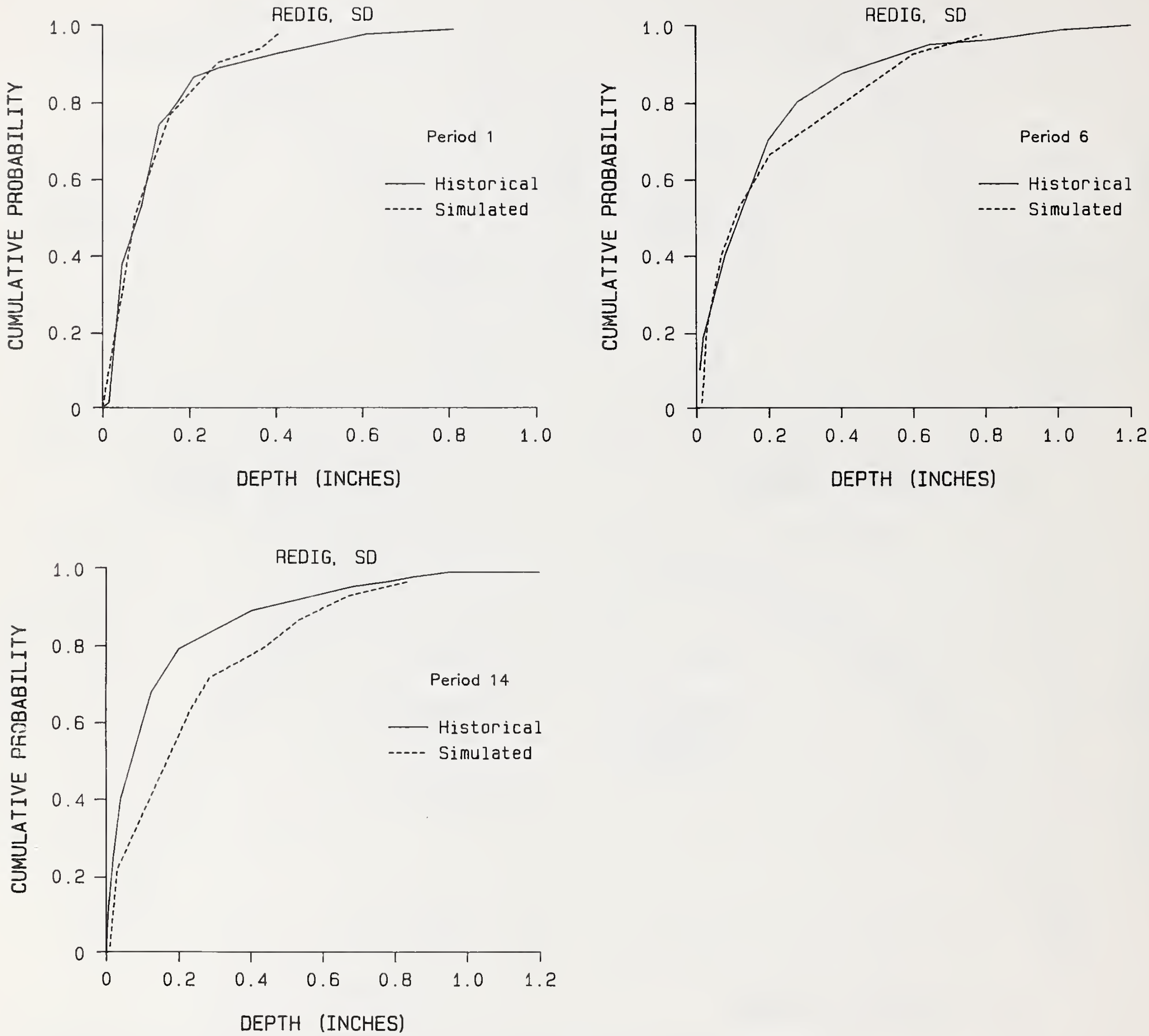
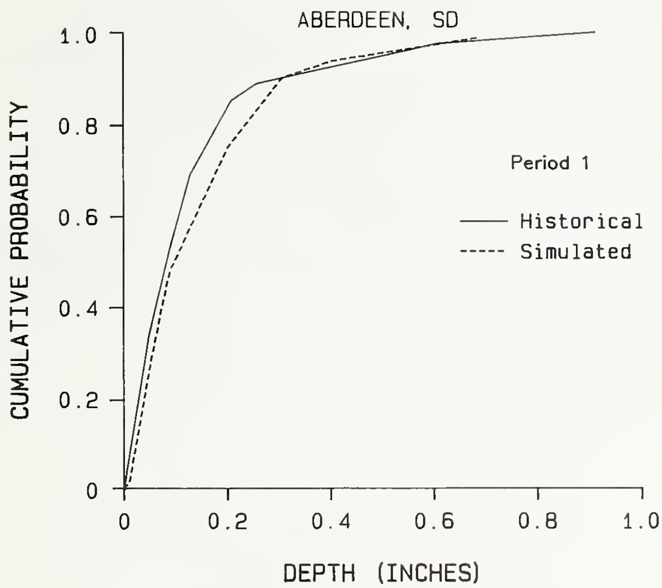
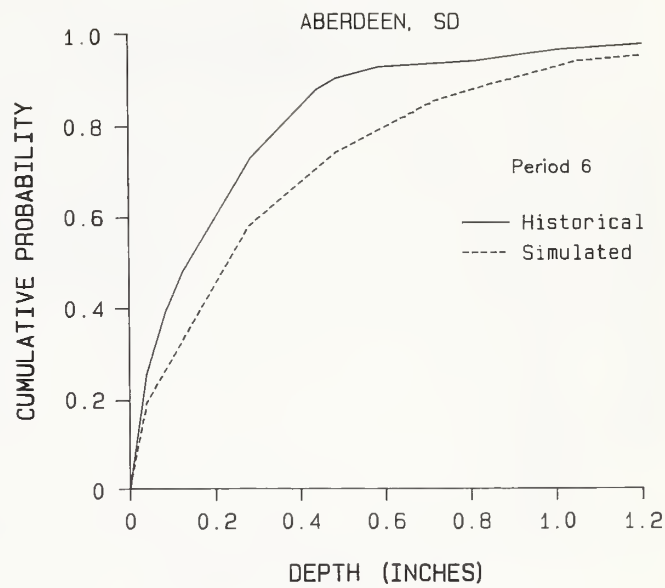
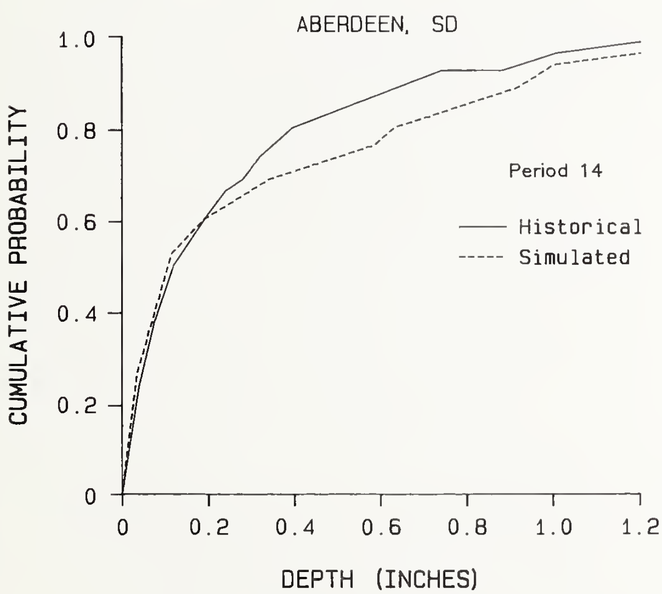


Figure 11.
Historical and simulated cumulative distribution functions of precipitation depth per wet day for three 14-day periods in Redig and Aberdeen, SD.


ABERDEEN, SD
Period 1
Historical
Simulated
CUMULATIVE PROBABILITY
DEPTH (INCHES)
0
0.2
0.4
0.6
0.8
1.0


ABERDEEN, SD
Period 6
Historical
Simulated
CUMULATIVE PROBABILITY
DEPTH (INCHES)
0
0.2
0.4
0.6
0.8
1.0
1.2


ABERDEEN, SD
Period 14
Historical
Simulated
CUMULATIVE PROBABILITY
DEPTH (INCHES)
0
0.2
0.4
0.6
0.8
1.0
1.2

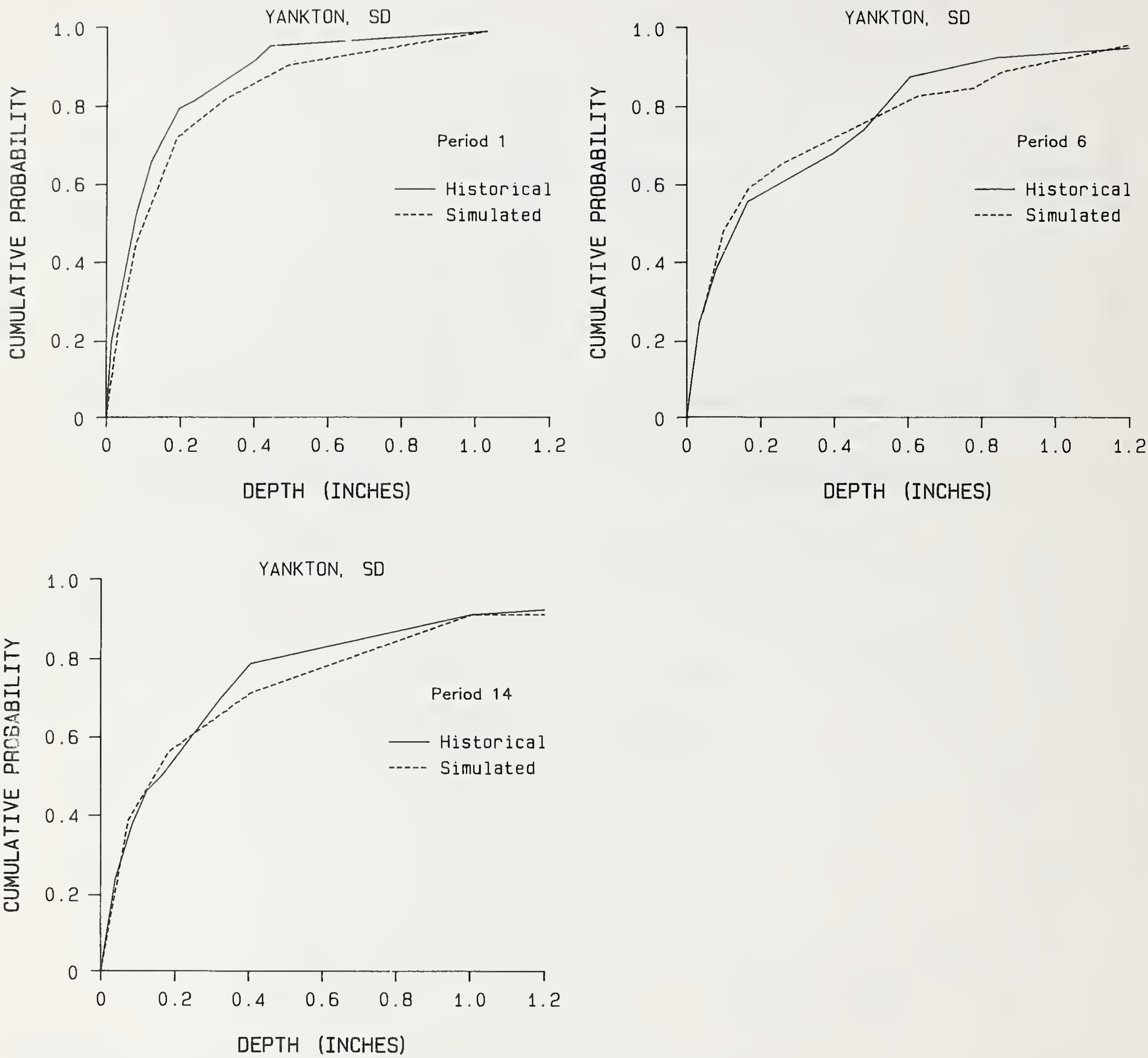
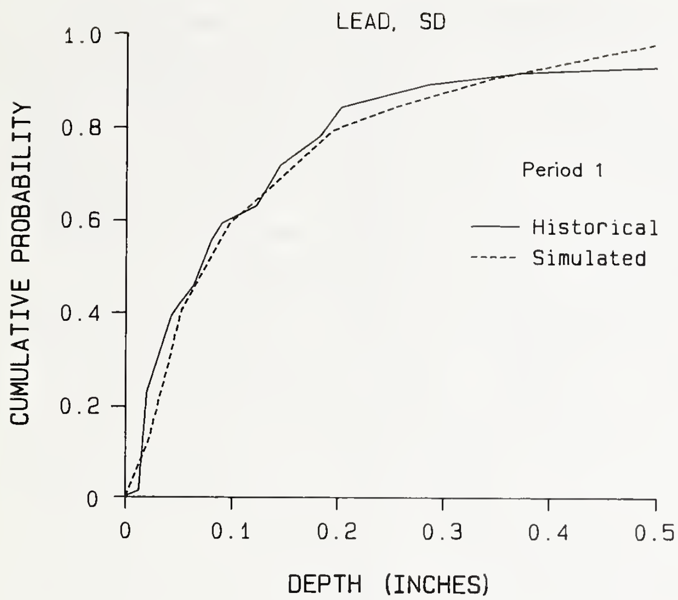
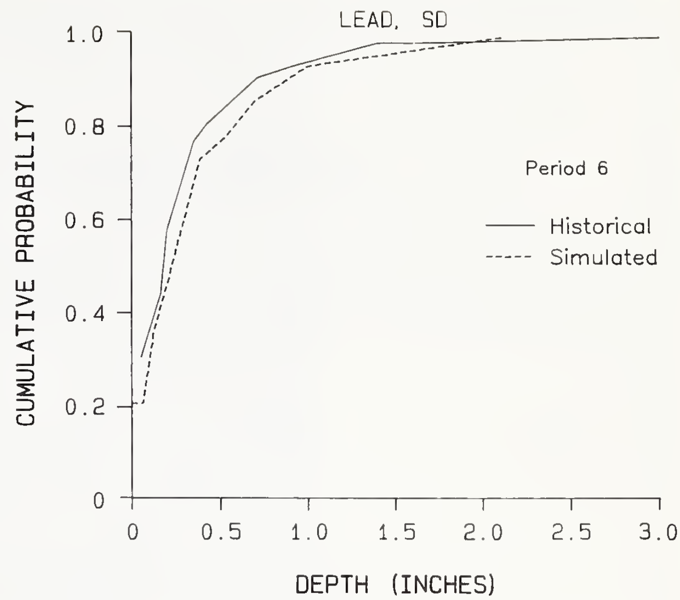
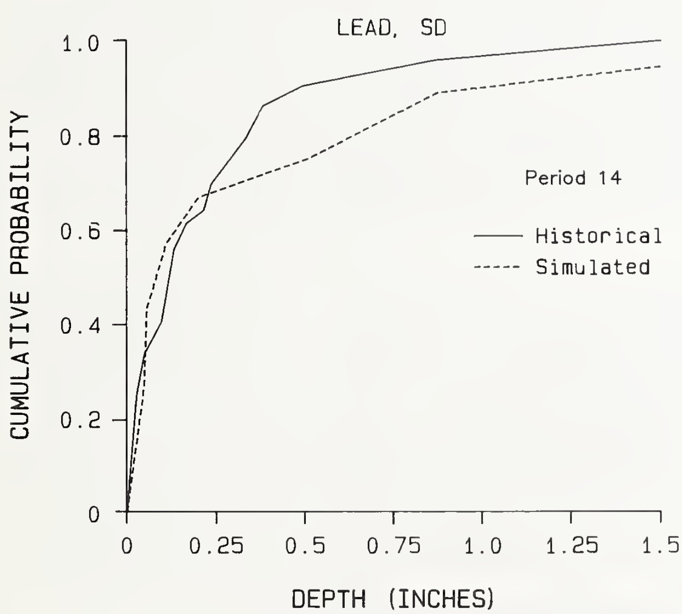


Figure 12.
Historical and simulated cumulative distribution functions of precipitation depth per wet day for three 14-day periods in Yankton and Lead, SD.


LEAD, SD
Period 1
Historical
Simulated
CUMULATIVE PROBABILITY
1.0
0.8
0.6
0.4
0.2
0
0
0.1
0.2
0.3
0.4
0.5
DEPTH (INCHES)


LEAD, SD
Period 6
Historical
Simulated
CUMULATIVE PROBABILITY
1.0
0.8
0.6
0.4
0.2
0
0
0.5
1.0
1.5
2.0
2.5
3.0
DEPTH (INCHES)


LEAD, SD
Period 14
Historical
Simulated
CUMULATIVE PROBABILITY
1.0
0.8
0.6
0.4
0.2
0
0
0.25
0.5
0.75
1.0
1.25
1.5
DEPTH (INCHES)

EXAMPLE

To run the program, load BASICA in the IBM-PC or compatible computers or ZBASIC for the Z-100. Then place the program disk in drive B and load "B:CLIMATPC.BAS" for the IBM-PC and compatibles or "B:CLIMATZ.BAS" for the Z-100 and give the RUN command. The welcome screen (fig. 13) will then appear. Press RETURN for the next screen (fig. 14) which provides two options. In this example we will assume that we wish to provide climatic information for a location west of Brookings and that we have not stored the parameters so we will elect option A. Therefore, we enter A <R>. Screen 3 then follows (fig. 15) explaining what the following screen will be. After pressing RETURN a map of the State appears (fig. 16) showing the locations of the weather stations. After the precipitation parameter files are read for each station, the cursor will appear in the lower left corner of the screen. We will use the arrow keys to move the cursor to the desired location and then press RETURN. Two circles will appear centered on the point of interest (fig. 17). The small circle has a radius of 30 miles and the larger a radius of 100 miles. If there is a station within the 30-mile radius, you will be asked if you wish to use parameters for that station only. If you answer yes, Y, you will have a nearest neighbor estimate of parameters; if you answer no, N, you will be asked if you wish to eliminate any of the stations within the 100-mile radius. In this case we will assume that no stations are to be omitted; therefore, we respond with N <R>. Next you will be asked if you wish to store parameters. If you are planning to use information for this location frequently, respond with Y and follow instructions for naming the file. When you enter the program again, select option B above. For this example we will respond with N.

At this point the program will calculate the daily precipitation model parameter values (which will take from 1 to 2 minutes) and then will display on the screen the averaged means, amplitudes, and phase angles for each parameter (fig. 18a). These values may be useful for diagnostic purposes. The program will then compute the theoretical mean annual precipitation using equation (10) and will ask if you know the mean annual precipitation. If the response is yes, you will be requested to enter the value in the form xx.xx (inches). The program will then adjust the parameters $\bar{p}_{10}$ and $\bar{\alpha}$ until the theoretical annual precipitation is within 0.1% of the known value (see Fig. 18b).

WELCOME TO SOUTH DAKOTA WEATHER

This program provides easy access to information on daily precipitation, maximum and minimum temperature, and solar radiation.

Press RETURN to continue.

Figure 13.
Welcome screen for microcomputer program.

THE FOLLOWING OPTIONS ARE AVAILABLE.

A. Estimate daily weather parameters. Use this option if this is the first time you have used this program for the location of interest.

B. Read daily weather parameters from a disk file. This option is used if you have already used option A and stored the parameters.

ENTER LETTER CORRESPONDING TO OPTION YOU DESIRE.?

Figure 14.
Option screen

The next display will be a map of the state of South Dakota

with locations of the weather stations where the data were obtained.

USE THE ARROWS ON YOUR KEYBOARD TO PLACE THE CURSOR AT THE POSITION

WHERE INFORMATION IS DESIRED. THEN PRESS RETURN TO CONTINUE.

Figure 15.
Information screen

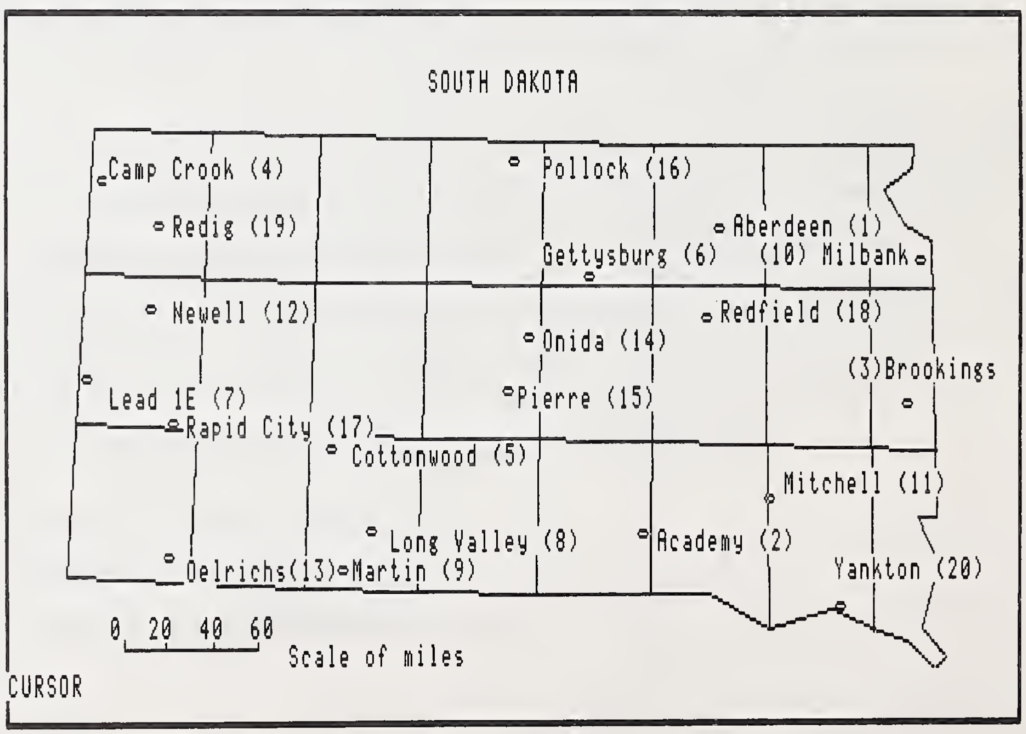


Figure 16.
Map of State with location of weather stations.

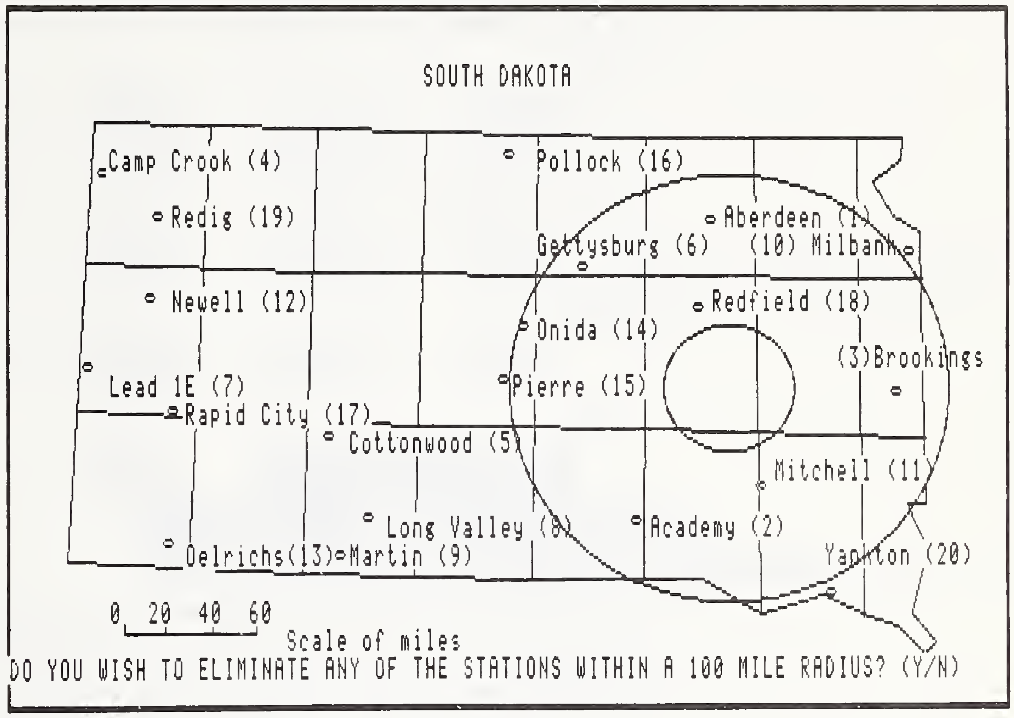


Figure 17.
Map of State with concentric circles around location where weather data are desired.

A

```
AVERAGE PARAMETERS FOR STATION

P00
   0.8354E+00     0.7611E-01     0.2799E+01     0.1366E-01     0.8844E+00
   0.0000E+00     0.0000E+00     0.6292E-02     -.1643E+01     0.2742E-02
   0.1225E+01     0.0000E+00     0.0000E+00
P10
   0.6639E+00     0.7735E-01     -.3033E+01     0.6411E-01     0.2523E+01
   0.1325E-01     0.3261E+00     0.0000E+00     0.0000E+00     0.3733E-02
   0.1667E+01     0.5997E-02     -.1988E+00
BETA
   0.7580E-01     0.2537E-01     -.6008E+00     0.0000E+00     0.0000E+00
   0.2914E-02     -.2842E+01     0.0000E+00     0.0000E+00     0.7795E-03
   -.6608E+00     -.1440E-02     0.1218E+01
MEAN
   0.2370E+00     0.1183E+00     -.7571E+00     0.5248E-02     0.3016E+00
   0.0000E+00     0.0000E+00     0.3780E-02     -.1155E+01     0.0000E+00
   0.0000E+00     0.0000E+00     0.0000E+00
ALPHA =  .4129134

PRESS RETURN TO CONTINUE?
```

B

```
CALCULATING
EXPECTED ANNUAL PRECIPITATION = 19.63667 INCHES. NO. WET DAYS= 72.6722

DO YOU KNOW THE AVERAGE ANNUAL PRECIPITATION? (Y/N)  Y

ENTER AVERAGE ANNUAL PRECIPITATION IN INCHES XX.XX? 20.20

CALCULATING
EXPECTED ANNUAL PRECIPITATION = 20.29476 INCHES. NO. WET DAYS= 73.78923

CALCULATING
EXPECTED ANNUAL PRECIPITATION = 20.18584 INCHES. NO. WET DAYS= 73.60731

ALPHA AND P10 HAVE BEEN ADJUSTED TO MAINTAIN ANNUAL PRECIP.

PRESS RETURN TO CONTINUE?
```

C

```
THE FOLLOWING OPTIONS ARE AVAILABLE

  A. Estimate the probability of X or less inches of rain in N days.

  B. Simulate M years of rainfall data, or rainfall, max

     and min temperature and radiation data.

ENTER LETTER OF OPTION DESIRED OR Q TO QUIT?
```

Figure 18.
Screen displays: a, parameters for station selected; b, iterative procedure to adjust average annual precipitation by modifying α and P_{10}; c, options.

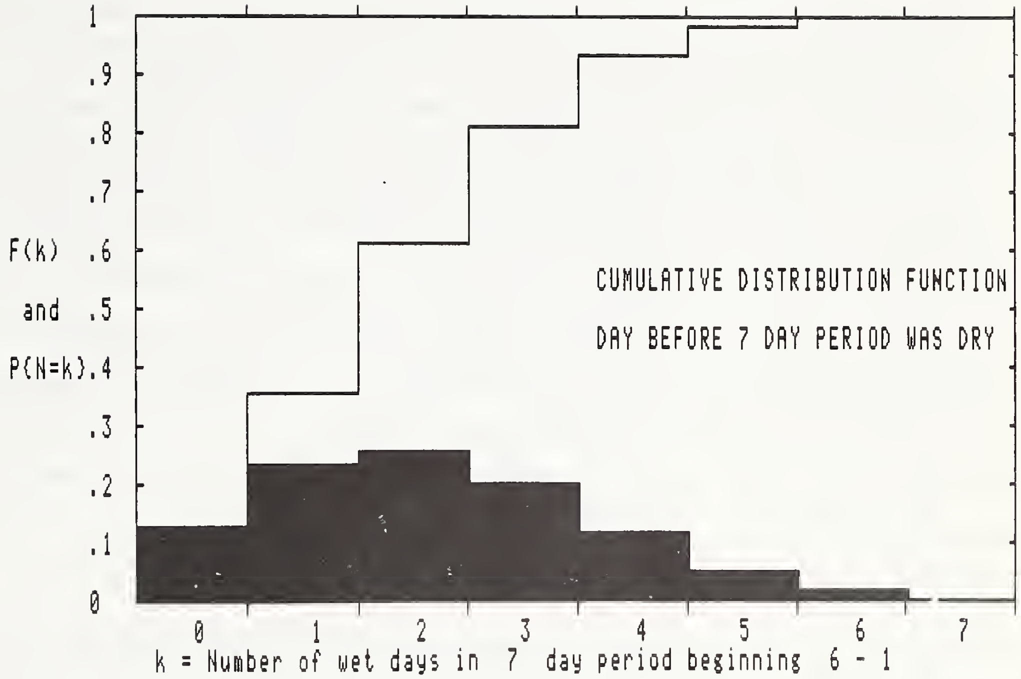


Figure 19.
Cumulative distribution function and probability mass function of the number of wet days in a 7-day period.

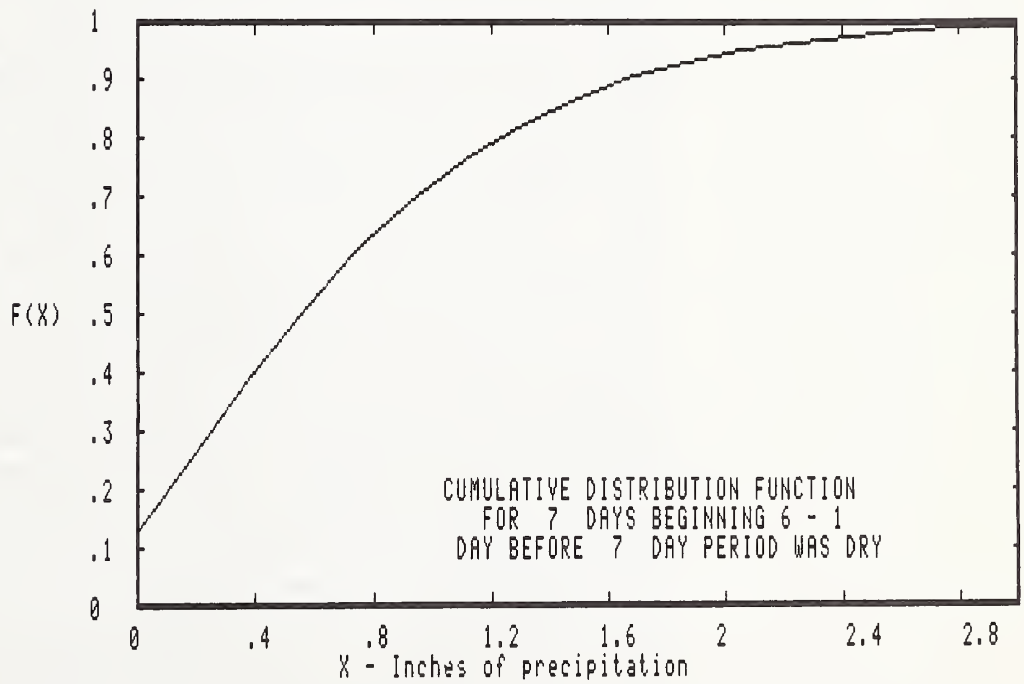


Figure 20.
Cumulative distribution function of total precipitation in a 7-day period.

```
This is the daily rainfall simulation option.
You will be asked how many years of data you wish to simulate
and the name of the sequential file to store the data in.
Each year of precipitation, maximum and minimum temperature
and radiation data requires about 9,150 bytes.  Thus the maximum
number of years on a 320 K-byte disk is about 30.
For precipitation only, about 100 years is acceptable.
PRESS RETURN TO CONTINUE?
```

Figure 21.
Information screen for simulation option.

Next, the screen will provide options to calculate probabilities of various amounts of rainfall for an m-day (m<30) period, beginning on any day of the year, or to simulate n years of climatic data (Fig. 18c). If option A, calculation of rainfall probabilities, is chosen, the user will be prompted to select the beginning day of the period, the number of days in the period, and finally will be asked if the day before the m day period was "dry," "wet," or "don't know." If you wish rainfall probabilities for the next m days enter today's state (wet or dry). If, on the other hand, you wish rainfall probabilities for a day more than one day in the future enter "DK." If your response is "DK," you will be asked if you have an estimate of the probability of rain on the day before. If you respond "N," the unconditional probability of a wet day on the day before is used. If you respond "Y," you are asked to enter the probability of rain. This option might be used if, for example, it was June 1 and you were interested in the probability of receiving 1 inch or less rainfall in the 5 day period June 3 - June 7. If the local weather forecast gave a 20 percent chance of rain on June 2 you could enter 0.20 at the query. The probabilities of 0, 1, 2,... M wet days will be calculated using equations (14) and (15) conditioned on the state of the day before the period begins. Figure 19 shows the cumulative probabilities (the stair step function) of K or fewer wet days and the probability of exactly k wet days in the period. From this figure we can see that the most probable number of wet days is 2 and that the probability of 0, 1, or 2 wet days is approximately 0.6. The probability of no wet days is about 0.13. From figure 20, we see that the probability of no rain (which corresponds to the probability of no wet days) is about 0.13. The chances are about 9 out of 10 (probability of 0.9) of receiving 1.6 inches or less rain. Thus the chances are 1 out of 10 of receiving over 1.6 inches of rain.

At this point we can go back to the menu (fig. 18c) and elect the precipitation probability option again or the simulation option. If the simulation option (B) is chosen, a sequence of instructions appears on the screen (fig. 21). By providing the proper input at the prompts, the user will obtain a sequential file of M years of simulated precipitation data or precipitation, maximum temperature, minimum temperature, and radiation data on a disk. These data will be in the sequence $Y_1(1)$, $t_{max1}(1)$, $t_{min1}(1)$, $r_1(1)$, $Y_1(2)$, $t_{max1}(2)$, $t_{min1}(2)$, $r_1(2) \cdot \cdot \cdot Y_1(365)$, $t_{max1}(365)$, $t_{min1}(365)$, $r_1(365)$, $Y_2(1) \cdot \cdot \cdot Y_M(365)$, $t_{maxM}(365)$, $t_{minM}(365)$, $r_M(365)$ and may be used as input to other programs. Note that $Y_\tau(1)$ is the precipitation for March 1.

DISCUSSION

The weather information provided in CLIMATE.BAS is most useful in conjunction with additional programs that require daily precipitation only or precipitation, maximum and minimum temperatures, and radiation as input. For example, daily precipitation only could be used as input to a program providing estimates of daily runoff using the SCS curve number method. Several sequences of simulated precipitation, maximum and minimum temperature, and radiation for short periods (weeks) could be used in conjunction with models of plant growth, nitrogen uptake, leaching, and transformations to assist in short-term farm management of nitrogen fertilizer application to reduce N losses from the root zone. Such sequences could also assist in estimating trafficability in military, construction, or agricultural operations. Sequences of annual simulations could serve as input for models evaluating chemical transport (Knisel 1980), soil erosion, and plant growth (Jones and Kiniry 1986). The analytic calculations of probabilities of M-day precipitation provide quick estimates of risk in weather-dependent activities.

REFERENCES

Akaike, H. 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control AC-19(6):716-723.

Dethier, B.E., and J.K. McGuire. 1961. The climate of the Northeast. Probability of selected weekly precipitation amounts in the Northeast region of the U.S., Cornell Univ. Agric. Exp. Stn. Agronomy Mimeo., 61-64 pp.

Feyerherm, A.M., and L.D. Bark. 1965. Statistical methods for persistent precipitation patterns. J. Appl. Meteorol. 4:320-328.

Feyerherm, A.M., L.D. Bark, and W.C. Burrows. 1965. Probabilities of sequences of wet and dry days in South Dakota. North Central Region Res. Pub. 161 and Tech. Bull. 139h. Kans. Agric. Exp. Stn., Manhattan, KS. 55 pp.

Gabriel, K.R. 1959. The distribution of the number of successes in a sequence of dependent trials. Biometrika 96:959-960.

Gabriel, K.R., and J. Neuman. 1962. A Markov chain model for daily rainfall in Tel Aviv. Q. J. R. Meteor. Soc. 88:90-95.

Gifford, R.O., G.L. Ashcroft, and M.D. Magnuson. 1967. Probability of selected precipitation amounts in the Western Region of the United States, Western Regional Research Publication and T-8 Nevada Agric. Exp. Stn., Reno, NV.

Jones, C.A., and J.R. Kiniry (eds.). 1986. CERES-Maize. A simulation model of maize growth and development. Texas A&M University Press, College Station, TX. 194 pp.

Knisel, W.G. (ed.) 1980. CREAMS - A field scale model for chemical runoff and erosion from agricultural management systems. USDA-SEA Conservation Research Report No. 26. 643 pp.

Lane, L.J. 1984. Surface water management: A users guide to calculate a water balance using the CREAMS model. LA-10177-M Manual UC-70B. Los Alamos National Laboratory, Los Alamos, NM. 49 pp.

Matalas, N.C. 1967. Mathematical assessment of synthetic hydrology. Water Resour. Res. 3(4):937-945.

Naylor, T.H., J.L. Balintfy, D.S. Burdick, and K. Chu. 1966. Computer simulation techniques. John Wiley & Sons, New York, 352 pp.

Nelder, J.A. and R. Mead. 1965. A simplex method for function minimization. Compu. J. 7(4):308-313.

Osborn, H.B., C.L. Hanson, and D.A. Woolhiser. 1987. Effect of elevation and aspect on daily precipitation model parameters. In Proceedings 7th Conference on Hydrometeorology, Edmonton, Alberta, Canada. American Meteorological Society.

Palmer, W.L., B.J. Barfield, and C.T. Haan. 1982. Sizing farm reservoirs for supplemental irrigation of corn. Part I: Modeling reservoir size yield relationships. Trans. ASAE 25:372-376.

Richardson, C.W. 1981. Stochastic simulation of daily precipitation, temperature, and solar radiation. Water Resour. Res. 17(1):182-190.

Richardson, C.W. 1982. Dependence structure of daily temperature and solar radiation. Trans. ASAE 25(3):735-739.

Richardson, C.W., and D.A. Wright. 1984. WGEN: A model for generating daily weather variables. U.S. Department of Agriculture, Agricultural Research Service ARS-8, 83 pp.

Roldan, J., and D.A. Woolhiser. 1982. Stochastic daily precipitation models. 1. A comparison of occurrence processes. Water Resour. Res. 18(5):1451-1459.

Shaw, R.H., G.L. Barger, and R.F. Dale. 1960. Precipitation probabilities in the North Central States. Missouri Agric. Exp. Stn. Bull. 753, Columbia, MO. 72 pp.

Smith, R.E., and H.A. Schreiber. 1974. Point processes of seasonal thunderstorm rainfall. 2. Rainfall depth probabilities. Water Resour. Res. 10(3):418-423.

Stern, R.D., and R. Coe. 1984. A model fitting analysis of daily rainfall data. J. R. Statist. Soc. A, 147, Part 1, pp. 1-34.

Todorovic, P., and D.A. Woolhiser. 1975. A stochastic model of n-day precipitation. J. Appl. Meteorol. 14(1):17-24.

Von Bargen, K. 1967. A systems approach to harvesting alfalfa hay. Trans. ASAE 10(3):318-319.

Woolhiser, D.A., and G.G.S. Pegram. 1979. Maximum likelihood estimation of Fourier coefficients to describe seasonal variation of parameters in stochastic daily precipitation models. J. Appl. Meteorol. 18(1):34-42.

Woolhiser, D.A.., and J. Roldan. 1986. Seasonal and regional variability of parameters for stochastic daily precipitation models - South Dakota, U.S.A. Water Resour. Res. 22(6):965-978.

Woolhiser, D.A., and J. Roldan. 1982. Stochastic daily precipitation models. 2. A comparison of distributions of amounts. Water Resour. Res. 18(5):1461-1468.

Zucchini, W. and P.T. Adamson. 1984. The occurrence and severity of droughts in South Africa. WRC Report No. 91/1/84, Dept. of Civil Engineering, Univ. of Stellenbosch and Dept. of Water Affairs. 198 pp.

APPENDIX: PARAMETER ESTIMATION FOR THE MARKOV-CHAIN/ MIXED-EXPONENTIAL MODEL

The maximum likelihood estimation of Fourier coefficients describing the seasonal variation of the Markov-chain/mixed exponential model parameters is accomplished by the FORTRAN program, AGUA46. The procedures used in the program are similar to those described by Woolhiser and Pegram (1979), Roldan and Woolhiser (1982), and Woolhiser and Roldan (1986). However, a few improvements have been made, so the general optimization strategy will be reviewed here. The discussion will parallel the order of computations in AGUA46.

AGUA46 consists of the main program, AGUA46 and 23 subroutines. The principal functions of each subroutine are discussed below and are also indicated in comments in the program on the diskette.

PROGRAM AGUA46: Reads input information and precipitation data and calls subroutines MARKOV, MARLIK, CALCUL and PLOT.

Subroutines

MARKOV: Called from AGUA46. Calculates the transition probabilities $p_{00}(k)$ and $p_{10}(k)$, k=1, ... NPER, for each period of the year as defined by the input parameter, NPER. Currently NPER may be 26 for 14-day periods or 13 for 28-day periods. The first period starts on March 1 and the last period contains 15 or 29 days. The extra day in leap year is ignored. The maximum likelihood parameter estimates are

$$p_{00}(k) = \frac{a_{00}(k)}{a_{00}(k) + a_{01}(k)} \tag{A.1}$$

$$p_{10}(k)= \frac{a_{10}(k)}{a_{10}(k) + a_{11}(k)} \tag{A.2}$$

where $a_{ij}(k)$ is the observed number of transitions from state i to state j in period k for the entire period of record.

FOUTER: Called from MARKOV and CALCUL. Calculates the least squares Fourier coefficients for the MC-ME parameters, p_{00}, p_{10}, α, β, δ and the mixed exponential mean daily depth, μ. The Fourier series is fit to the parameter values at the period midpoints. Calls subroutine SERIES.

SERIES: Called from FOUTER. Calculates the least squares Fourier coefficients. Calls subroutines COEFF and FOURI.

COEFF: Called from SERIES. Computes least squares Fourier coefficients using a trapezoidal rule to approximate the integral.

FOURI: Called from SERIES. Computes values of the Fourier series at period midpoints given the coefficients.

MARLIK: Called from program AGUA46. Calculates approximate ML estimators for the amplitudes and phase angles for the Fourier series approximation to the parameters $p_{00}(n)$ and $p_{10}(n)$ by maximizing the following functions:

Dry-dry transitions:

$$\text{Log } L_d = \sum_{n=1}^{365} \{a_{00}(n)\log p_{00}(n) + a_{01}(n)\log[1 - p_{00}(n)]\} \quad \text{(A.3)}$$

Wet-dry transitions:

$$\text{Log } L_w = \sum_{n=1}^{365} \{a_{10}(n)\log p_{10}(n) + a_{11}(n)\log[1-p_{10}(n)]\} \quad \text{(A.4)}$$

where

$$p_{i0}(n) = \bar{p}_{i0} + \sum_{j=1}^{m_{i0}} \{C_{i0j} \sin(2\pi nj/365 + \phi_{i0j})\} \quad \text{(A.5)}$$

$a_{ij}(n)$ = observed number of transitions from state i on day n-1 to state j on day n, where i = 0,1, m_{i0} is the maximum number of harmonics to be considered, C_{i0j} is the amplitude of the jth harmonic and ϕ_{i0j} is the phase angle of the jth harmonic. Calls subroutines DLIK and WLIK through SIMPLX.

DLIK: Called from MARLIK through SIMPLX. Calculates the log likelihood function for the transitions beginning with dry days, equation (A.3). A penalty function is added if any $p_{00}(n)$ is negative or greater than 1.

WLIK: Called from MARLIK through SIMPLX. Calculates the log likelihood function for the transitions beginning with wet days, equation (A.4). A penalty function is added if any $p_{10}(n)$ is negative or greater than 1.

CALCUL: Called from program AGUA46. This subroutine calculates the mean and variance of precipitation per wet day, the mean number of wet days, and the mean and variance of precipitation totals for each 14 or 28 day period. It also sums the annual precipitation for each year and calls the subroutine PLOT to plot the annual series on the line printer. The subroutine ESTADI is called to calculate the period statistics, and MIXEXP is called to calculate the parameters of the mixed exponential distribution for each period. The subroutine FOUTER is called to calculate the least squares Fourier coefficients to fit the series through the period values for the parameters α_k, β_k, and δ_k for the mixed exponential distribution. These coefficients

are used as starting values in the subroutine LIKMEX which calculates maximum likelihood estimates of the Fourier coefficients.

ESTADI: Called from CALCUL. Computes the mean and variance of daily precipitation amounts for 14 or 28 day periods.

MIXEXP: Called from CALCUL. Computes the parameters α_k, β_k, and δ_k for the mixed exponential distribution for daily precipitation for each 14- or 28-day period by maximizing the log likelihood:

$$\text{Log } L_k = \sum_{j=1}^{N(k)} \{\log[\alpha_k/\beta_k \exp(-u_{kj}/\beta_k) + (1-\alpha_k)/\delta_k \exp(-u_{kj}/\delta_k)]\} \qquad (A.6)$$

The sample values, u_{kj}, consist of the observed daily precipitation minus the threshold. Optimization is achieved through calls to the subroutine SIMPLX. The subroutine FPLIK is called through SIMPLX to evaluate the log likelihood function.

FPLIK: Called from MIXEXP through SIMPLX. Calculates the objective function equation (A.6) for each 14- or 28-day period. A penalty function is used to prevent the parameters from getting out of the appropriate range, that is,

$$0 \le \alpha_k \le 1, \quad 0 < \beta_k, \ 0 < \delta_k.$$

LIKMEX: Called from CALCUL. Calculates maximum likelihood estimates of Fourier series coefficients for the mixed exponential parameters $\alpha(n)$, $\beta(n)$, and the mean $\mu(n)$, where

$$\mu(n) = \alpha(n)\ \beta(n) + (1 - \alpha(n))\delta(n)$$

The log likelihood function to be maximized is

$$\text{Log}_{ME} = \sum_{n=1}^{365} \sum_{j=1}^{m(n)} \{\log[\alpha(n)/\beta(n)\exp[-u_{nj}/\beta(j)] + ((1-\alpha(n))/\delta(n)\exp[-u_{nj}/\delta(n)]]\} \qquad (A.7)$$

where m(n) = number of wet days on day n, n=1,2 ... 365 for the period of record., u_{nj} = the transformed precipitation for the jth wet day of day n. The parameters $\alpha(n)$, $\beta(n)$, and $\mu(n)$ are represented in the form of equation (A.5).

Three options are available for representing the parameter $\alpha(n)$.

1. $\alpha(n)$ is a constant throughout the year and is provided as input to the program. This case might be used if a single value

of α is to be used for a given region. For this option CALFA > 0 and KALFA > 0.

2. $\alpha(n)$ is a constant throughout the year and is estimated by taking the mean of the α_k values for each period. CALFA < 0 and KALFA > 0. <u>This option is not recommended</u>.

3. $\alpha(n)$ is fit by Fourier series with maximum number of harmonics, MAXA. CALFA < 0 and KALFA < 0. This is the recommended option for most circumstances.

These options were included to provide flexibility in the analysis. Through experience we have found that the parameters α_k are quite variable and that there are strong interactions between α_k, β_k, and δ_k. By setting α as a constant (MAXA = 0) throughout the year, problems with interactions are reduced.

The optimization strategy is described below.

1. The mean values of the parameters, $\overline{\alpha}$, $\overline{\beta}$, $\overline{\mu}$, are first estimated simultaneously by calling subroutine MDLIK through SIMPLX.

2. The amplitude and phase angle of the first harmonic of the mean, μ, are estimated by calling subroutine MULIK through SIMPLX.

3. The amplitude and phase angle of the first harmonic of β are estimated by calling subroutine BLIK through SIMPLX.

4. If α is to vary seasonally, the amplitude and phase angle of the first harmonic are estimated by calling AMLIK through SIMPLX.

5. Steps 2 and 3 or steps 2, 3 and 4 are repeated with 2nd and higher harmonics until the maximum number of harmonics, MAXH, has been reached. As the parameters for each harmonic have been estimated by the maximum likelihood technique, a decision is made to retain or drop that harmonic depending on the value of the Akaike Information Criterion, AIC (Akaike 1974),

$$AIC = -2(\mathrm{Log}_{ME} - n) \tag{A.8}$$

where n is the current number of parameters. If the value of AIC is smaller than the previous one (that is, before the harmonic was added) the harmonic is assumed to be significant. If AIC is greater than the previous value the harmonic is assumed to be insignificant and is dropped from consideration.

6. A second round of optimization is started by obtaining improved estimates of the mean parameters $\overline{\alpha}$, $\overline{\beta}$, and $\overline{\mu}$ by calling MDL2K through SIMPLX.

7. Improved estimates of the amplitudes and phase angles of the mean $\mu(n)$ are estimated for each significant harmonic by repeatedly calling MULIK2 through SIMPLX.

8. Improved estimates of the amplitude and phase angles of the parameter $\beta(n)$ are estimated for each significant harmonic by repeatedly calling BLIK2 through SIMPLX.

MLIK: Called from LIKMEX, AMLIK, BLIK, MULIK, MDLIK, MDL2K, MULIK2 and BLIK2. This subroutine calculates the value of the objective function specified by equation (A.7)

AMLIK: Called from LIKMEX through SIMPLX. Called when the amplitude and the phase angle of a harmonic of $\alpha(n)$ are being optimized. Calls subroutine MLIK.

BLIK: Called from LIKMEX through SIMPLX: Called when an amplitude and a phase angle of a harmonic of $\beta(n)$ are being optimized. Calls subroutine MLIK.

MULIK: Called from LIKMEX through SIMPLX. Called when an amplitude and a phase angle of $\mu(n)$ are being optimized. Calls subroutine MLIK.

MDLIK: Called from LIKMEX through SIMPLX. Called when the means of the three parameters α, β, and μ are being optimized for the first time. Calls subroutine MLIK.

MDL2K: Called from LIKMEX through SIMPLX. Called when the means of the three parameters α, β, and μ are being optimized for the second time. Amplitudes and phase angles of significant harmonics of $\alpha(n)$, $\beta(n)$ and $\mu(n)$ are retained. Calls subroutine MLIK.

PLOT: Called from AGUA46. Plots data on the line printer.

MULIK2: Called from LIKMEX through SIMPLX. Called when an amplitude and a phase angle of $\mu(n)$ are being optimized for the second time. Calls subroutine MLIK.

BLIK2: Called from LIKMEX through SIMPLX. Called when an amplitude or phase angle of $\beta(n)$ are being optimized for the second time. Calls subroutine MLIK.

SIMPLX: Called from MARLIK and LIKMEX. This is an unconstrained multivariate optimization routine based on work by Nelder and Mead (1965).

The version of AGUA46 included on the disk has been run on a VAX 11/750 computer. The program may need modifications for other computers. A run for Aberdeen, SD required 20 min. of CPU time on the VAX 11/750. The techniques used by Stern and Coe (1984) would require considerably less computer time. However, it would be necessary to replace the mixed exponential distribution with the gamma distribution and to extensively revise the microcomputer program.